DOCUMENT RESUME

ED 062 309 SP 005 706

AUTHOR	Lopossa, Barbara D.
TITLE	A Comparative Study of Team and Individual Decision Making.
INSTITUTION	Stanford Univ., Calif. Stanford Center for Research and Development in Teaching.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	TR-20
BUREAU NO	BR-5-0252
PUB DATE	Sep 71
CONTRACT	OEC-6-10-078
NOTE	73p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Cognitive Processes; *Decision Making; *Decision Making Skills; *Leadership Responsibility; *Teacher Responsibility

ABSTRACT

The aims of this study were a) to compare the quality of decisions and the decision-making behavior of individual teachers and teaching teams, and b) to study the effects of formal leadership on decision making. One hundred and ninety teachers from California elementary schools were divided and placed into sections of teaching teams or self-contained classrooms. Of the teaching team section, 75 teachers were assigned to work in 20 teams, and 20 teachers were assigned to work as individuals. The other group from self-contained classrooms included 75 teachers assigned to 20 ad hoc groups and 20 working as individuals. Decision-making behavior was measured by the way in which subjects evaluated a given set of alternative courses of action in regard to each of two problems of student behavior. The decision consisted of ranking these alternatives after they had been evaluated. The quality of the decision was determined by comparing the subjects' ranking with an average ranking obtained from 15 experts. The verbal and nonverbal behavior of groups was recorded and analyzed using the Bales Interaction Process Analysis. Results showed that groups were more extreme than individuals in evaluating the consequences of teacher behavior, ad hoc groups displayed more tension than teams, and smaller groups showed more solidarity and less disagreement than did larger ones. A 25-item bibliography and appendixes are included. (Author/MJM)

ED 062309

STANFORD CENTER
FOR RESEARCH AND DEVELOPMENT
IN TEACHING

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Technical Report No. 20

A COMPARATIVE STUDY OF TEAM AND INDIVIDUAL DECISION MAKING

Barbara D. Lopossa

School of Education
Stanford University
Stanford, California

September 1971

Published by the Stanford Center for Research and Development in Teaching, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred. (Contract No. OEC-6-10-078, Project No. 5-0252-0307.)

SP005706

Introductory Statement

The Center is concerned with the shortcomings of teaching in American schools: the ineffectiveness of many American teachers in promoting achievement of higher cognitive objectives, in engaging their students in the tasks of school learning, and, especially, in serving the needs of students from low-income areas. Of equal concern is the inadequacy of American schools as environments fostering the teachers' own motivations, skills, and professionalism.

The Center employs the resources of the behavioral sciences--theoretical and methodological--in seeking and applying knowledge basic to the achievement of its objectives. Analysis of the Center's problem area has resulted in three programs: Heuristic Teaching, Teaching Students from Low-Income Areas, and the Environment for Teaching. Drawing primarily upon psychology and sociology, and also upon economics, political science, and anthropology, the Center has formulated integrated programs of research, development, demonstration, and dissemination in these three areas. In the Heuristic Teaching program, the strategy is to develop a model teacher training system integrating components that dependably enhance teaching skill. In the program on Teaching Students from Low-Income Areas, the strategy is to develop materials and procedures for engaging and motivating such students and their teachers. In the program on Environment for Teaching, the strategy is to develop patterns of school organization and teacher evaluation that will help teachers function more professionally, at higher levels of morale and commitment.

The following experimental study compares the rationality and quality of decision making by individuals and teams, by experienced teams and ad hoc groups, and by teams with leaders and teams without leaders. This early study should be read in light of two later reports from the Environment for Teaching program: Meyer et al., The impact of the open-space school upon teacher influence and autonomy (Technical Report No. 21); and Molnar, Teachers in Teams: Interaction, Influence, and Autonomy (Technical Report No. 22).

Contents

List of Tables v

List of Figures vi

Abstract vii

Introduction 1

Experimental Design 4

- Selection of Participating Districts and Subjects . . . 4
- Similarities of the Experimental Groups 5
- Description of Experimental Groups 6
- The Decision Process Test 8
- The Hypotheses 13
- Group Interaction 14
- Methods of Data Analysis 16

The Decision Process Test: The Results of Data Analysis . . 17

- Hypothesis 1. Groups Compared to Individuals 17
- Hypothesis 2. Teams Compared to Ad Hoc Groups 23
- Hypothesis 3. Team Teachers Compared to Self-Contained-Classroom Teachers 26
- Hypothesis 4. Teams with Leaders Compared to Teams without Leaders 26
- The Validity and Reliability of the Decision Process Test 26

Group Interaction: The Results of Data Analysis 32

- Distribution of Group Behavior: Pooled Interaction Profiles 32
- Group Participation Patterns 39
- Two-Way Analysis of Variance on Group Interaction . . . 42
- One-Way Analysis of Variance on Group Interaction . . . 44
- Significance of Interaction Data to Results on the Decision Process Test 45
- Interobserver Reliability 47

Summary of Findings 48

Discussion 50

- Sampling Procedures 51
- Procedures for Administering the Test 51
- Validity of the Decision Process Test 51

Conclusions 53

References 57

Appendix A: Decision Process Test 61

Appendix B: Expert Rating Sheet 65

List of Tables

1. Experimental Design 4
2. Characteristics of the Sample 7
3. Scoring System for Variables of Probability, Desirability, and Time Occurrence of Consequences 11
4. One-Way Analysis of Variance: Decision-Behavior Variables 18
5. One-Way Analysis of Variance: Decision-Quality Variables 22
6. Two-Way Analysis of Variance for Groups by Size and Type: Decision-Behavior Variables 24
7. Means and Standard Deviations for Groups by Size and Type: Decision-Behavior Variables 25
8. One-Way Analysis of Variance: Decision-Behavior Variables 27
9. Subjects' Ratings of Occurrence and Seriousness of Problems 28
10. Teacher Behaviors Subjects Would Actually Enact (number of times listed) 29
11. Decision Behavior: Correlation Between Scores on Problems IB and PA 30
12. Decision Quality: Correlation Between Scores on Problems IB and PA 31
13. Group Interaction: Two-Way Analysis of Variance for Groups by Size and Type 43
14. Group Interaction: Means and Standard Deviations for Groups by Size and Type 44
15. Group Interaction: One-Way Analysis of Variance . . . 45
16. Effects of High vs Low Socio-Emotional Behavior on Rationality and Quality of Decision 46
17. Number of Correct and Incorrect Predictions Based on Expected Utility Scores 53

B-1. Average Rank Given to Alternatives by Subgroups of Experts and Ss 67

B-2. Correlations Between Average Rank Given Alternatives for Subgroups of Experts and Total Group of Experts . . 68

List of Figures

1. Bales Interaction Process Categories 15
2. Pooled Interaction Profile: Teams 33
3. Pooled Interaction Profile: Ad Hoc Groups 34
4. Pooled Interaction Profile: Five-Member Groups 35
5. Pooled Interaction Profile: Three-Member Groups 36
6. Pooled Interaction Profile: Teams with Leaders 37
7. Pooled Interaction Profile: Teams without Leaders . . . 38
8. Distribution of Initiation of Interaction 40

Abstract

The aim of this study was to compare the quality of decisions and the decision-making behavior of individual teachers and teaching teams. In addition, experienced teams were compared with formative ones; and teams with appointed leaders were compared with teams without leaders in order to study the effects of formal leadership on decision making.

Decision-making behavior was measured by the way in which subjects evaluated a given set of alternative courses of action in regard to each of two problems of student behavior. The decision consisted of ranking these alternatives after they had been evaluated. The quality of the decision was determined by comparing the subjects' ranking with an average ranking obtained from 15 experts. The verbal and nonverbal behavior of groups was recorded and analyzed using the Bales Interaction Process Analysis.

190 teachers from California elementary schools participated. 95 of the teachers were from teaching teams; 75 of them were assigned to work in 20 teams and 20 were assigned to work as individuals. The other 95 subjects were teachers from self-contained classrooms; 75 of them were assigned to 20 ad hoc groups, and 20 worked as individuals.

An analysis of variance showed that groups were more extreme than individuals in evaluating the consequences of teacher behavior but that the quality of the decisions they reached did not differ markedly from those of the individuals. Nor, in general, did ad hoc groups differ from experienced teams in their decision-making behavior; in the analysis of group interaction, however, ad hoc groups were shown to display more tension than teams. Smaller groups, both teams and ad hoc groups, showed more solidarity and less disagreement than did larger ones. Teams without appointed leaders showed more solidarity and less disagreement than teams with leaders. All subjects tended to take into account only short-range consequences of their possible behavior. They anticipated only consequences that were either very desirable or very undesirable, and those that had at least a 70 percent chance of occurring.

A COMPARATIVE STUDY OF TEAM AND INDIVIDUAL DECISION MAKING

Barbara D. Lopossa

INTRODUCTION

Many advantages have been claimed for teaching teams as a form of elementary school organization. Most of these suggest that a group of teachers will be able to institute desirable changes in their instructional program and will become more effective problem solvers, generally, if they work together in teams instead of as individuals. If these assumptions were true, one would expect to see the improvement reflected in the academic progress and/or socio-emotional adjustment of students who have had the benefit of this method of instruction. So far, studies that have attempted to compare student gains under teaching teams and individual teachers have not shown that teams have produced many significant differences (for example, Knox, 1965; Lambert et al., 1964). However, most of these studies did not establish that the behavior of teams as it related to students actually differed from that of the individual teachers with whom they were compared. This may be one reason why the studies have failed to lend support to the claims made for team teaching and suggests that teacher variables may need to be examined before the variables concerned directly with student outcomes.

Therefore, one purpose of this study was to investigate teacher behavior--specifically, the selection of an appropriate course of action for handling discipline problems with a given student--and to determine whether or not teachers can be more effective as group members than as individuals in solving such problems. Shaplin (Shaplin & Olds, 1964) has suggested that the empirical findings from social-psychological research on small groups can be used as a theoretical framework for the study of team teaching, since he feels that team teaching represents a special case of this general problem. In that literature, there is some support for the assumption that groups might be expected to

---

Barbara Lopossa is Assistant Professor of Education at San Jose State College. The research reported here was carried out while Dr. Lopossa was a Research Assistant at the Stanford Center for Research and Development in Teaching. An abbreviated version of this report was presented at the Annual Meeting of the American Educational Research Association, Minneapolis, Minnesota, 1970.

make higher quality decisions than individuals. The results of numerous studies comparing the quality of small-group performances to that of individual performances over the past fifty years have led to the conclusion that, in general, group problem solving is superior (Collins & Guetzkow, 1964; Lorge et al., 1958). Brim et al. (1962) also found that when groups and individuals were asked to evaluate several courses of action and then to rank these actions in terms of desirability, groups tended to be more rational, that is, they tended to rank the alternatives more as they might be expected to do given their prior evaluations.

The differences noted between group and individual performance have often been slight, however, and many factors--especially the interpersonal relations of group members--have been shown to have negative effects on the quality of group performance. In addition, because most of these studies have been conducted with ad hoc groups, the applicability of the findings to established teaching teams is questionable. Furthermore, most of the tasks used in small-group research have been unrelated to the kinds of problems actually handled by real work groups. It would seem, then, that although some research findings support the superiority of ad hoc groups over individuals in decision making, much research remains to be done before generalizations about the superiority of teaching teams in this area can be made. Therefore, a second purpose of this study was to explore the applicability of findings from small-group research to team teaching by using them as a basis for hypotheses regarding some of the outcomes of this study. In order to make the necessary comparisons between experienced and formative groups as well as between groups and individuals, both experienced teaching teams and ad hoc groups of teachers were included.

Another reason for including both teams and ad hoc groups was to ascertain whether the experience of working together for a period of time would improve the performance of teams. Hall and Williams (1966) compared established groups and ad hoc groups of business executives and found that the factor of previous group experience improved performance on the experimental task and also the ability of the group to utilize initial differences of opinion among group members. There is also a sizable body of data on the effects of task and/or interpersonal-relations training on performance within work groups with the data generally indicating that such training improves performance (e.g. Maier, 1963). Therefore, teams whose members have worked together for a period of time and who may have received some instruction in the possibilities and problems of a team structure might be expected to perform better than ad hoc groups. A third purpose of this study was to investigate the effects of experience and possible training on teams.

If differences in decision-making ability do exist between teams and ad hoc groups of teachers, they may stem from the experience of working in groups. But it is also possible that

differences may be due to the personal characteristics of group members. Little is known about the selection processes by which some teachers are placed on teams and others are not. If team teachers had a certain cluster of personal attributes and teachers from self-contained classrooms another, these different qualities might affect the results of both individual and group work. It is also possible that problem-solving skills learned while working in a team might transfer to individual problem solving. Therefore, a fourth purpose of this study was to compare the decision-making ability of teachers assigned to teams and teachers assigned to self-contained classrooms both as groups and as individuals.

A final purpose was to explore the effects of having or not having an assigned leader on the decision-making behavior of experienced teams. Although many variations on team teaching exist, two types of teams have developed as school districts have experimented with this form of school organization. One is the cooperative team in which leadership is assumed at various times by different group members, depending on the nature of the problem under consideration. The other type of team is hierarchical, with at least two levels of assigned responsibility: formal leader and team member. Evidence from small-group research can be cited to support both types of organization (Hoffman, 1965; Maier, 1967). Therefore, the question of which type of organization is likely to make the better decisions was again raised in this study because of its practical concern to school administrators contemplating the move to team teaching.

This study was undertaken primarily to compare the decision-making ability of teams of elementary teachers with that of individual self-contained-classroom teachers. However, as outlined above, a number of other considerations emerged from a review of literature pertinent to team teaching. These were:

1. Do groups of teachers necessarily make better decisions than individual teachers?
2. To what extent, if any, are generalizations from small-group research applicable to the work of teaching teams?
3. What are the effects of training and experience in group work on teams?
4. What are the similarities and differences in the decision-making behavior of teachers assigned to teams and teachers assigned to self-contained classrooms?
5. Can any effects of formal versus emergent leadership on the decision making of experienced teams be determined? As explained below, the experiment was designed to yield information on each of these questions.

EXPERIMENTAL DESIGN

As indicated above, the main purpose of this study was to compare the decisions and decision-making behavior of elementary teachers under conditions of group problem solving and individual problem solving. In the group condition were both experienced teaching teams and ad hoc groups formed of teachers from self-contained classrooms. In the individual condition were teachers who normally worked on teams as well as teachers from self-contained classrooms. Thus there were four experimental conditions in all. Interaction effects were also examined, in order to obtain information on all of the questions listed above. The experimental design is shown in Table 1.

TABLE 1

Experimental Design

Teaching Assignment	Problem-Solving Condition		Total Units
	Group	Individual	
Team	I. Teaching Teams (N=20)	III. Individuals from Teams (N=20)	N=40
Self-contained Classroom	II. Ad hoc Groups (N=20)	IV. Individuals from Self-contained Classrooms (N=20)	N=40
Totals	Groups (N=40)	Individuals (N=40)	N=80

Selection of Participating Districts and Subjects

Seven suburban school districts in central and southern California provided the subjects for this study. These districts were chosen because they had elementary schools whose faculties were organized into teaching teams as well as other schools with a self-contained-classroom organization. The selection was further restricted to districts having teams which met the following

criteria: (1) the teams provided the major instruction in several areas of the curriculum for one group of students or, at the least, in one or more related areas for more than one group of students; (2) they held regular planning sessions; (3) they had no fewer than three members and no more than five members; and (4) the team members had worked together for at least the year of the data gathering (1967-68).

Because a large number of districts meeting the above criteria could not be identified, a random selection of teams could not be made, and all teams in the participating districts were used. A random selection was possible for a corresponding number of schools in each district having a self-contained-classroom organization from which the ad hoc groups were drawn. Selection for ad hoc groups was by school rather than by individual teacher because of the logistical problem imposed by trying to form ad hoc groups from different buildings. Seven of the potential ad hoc groups were randomly selected for placement of these teachers in an individual condition. Subjects for the other individual condition were obtained by randomly selecting seven of the available teams and assigning the members of these teams to the individual condition. Where not all members were needed to make up the required twenty subjects for the individual condition, some were randomly omitted from the experiment.

Similarities of the Experimental Groups

Data were collected on the sex, age, number of years of teaching experience, and number of masters degrees held by the subjects. These questions were asked so as to determine whether each of the experimental groups might be considered a sample from the same population, and to obtain statistics that would make possible a comparison between the teachers in this sample and elementary school teachers in general. Such comparisons are necessary in order to interpret results on the Decision Process Test as well as to determine if the findings can be generalized beyond this sample.

In order to compare subjects on the above personal factors, they were classified by each of the above as well as by the type of experimental condition. The Chi-square statistic was used to test the hypothesis that the two characteristics (i.e., sex and experimental condition) are independent. Thus the proportion of men and women in each condition was compared with the expected number of each based upon proportions in the total sample. The same was done for the number of teachers holding a B.A. degree and the number holding an M.A. On the factor of age, subjects were classified according to four age groups (20-29, 30-39, 40-49, and 50-65), and the actual number of subjects falling in each category was compared with the number expected, also using the Chi-square statistic. A similar procedure was used for years of teaching

experience, the subgroups representing 1-3 years, 4-9 years, 10-19 years, and 20-35 years of experience.

The above tests showed the four experimental conditions were not statistically different on the above characteristics (.05). However, team teachers tended to be younger and to have fewer years of teaching experience than did the self-contained-classroom teachers. A larger percentage of women teachers from self-contained classrooms in the individual condition had masters degrees than did the women in other groups.

When compared to the national averages for the 1965-66 school year, this sample of elementary teachers appeared to contain a higher percentage of men, and to be somewhat younger and less experienced than elementary teachers in general (NEA, 1967a). They seemed to be nearly comparable in terms of the number of advanced degrees held. However, since the data for this study were collected during the 1967-68 school year, these comparisons are predicated on the notion that averages remained stable during the intervening two years.

Table 2 shows how subjects compared on personal characteristics as well as how they compared with elementary teachers across the nation.

Description of Experimental Groups

Teachers were assigned to one of the four experimental groups indicated in Table 1. The group composition is described below:

Group I, Group condition: Teams. Twenty teams composed of seventy-five teachers were placed in the group problem-solving condition so that each team could work as an intact group. These were teams that had been established and functioning for at least eight months. The size of the teams varied: there were ten teams of three members each; five teams of four members each; and five teams of five members each. Ten of these teams worked with primary children and ten with children in intermediate grades. Ten of the teams had a leader appointed by the school administrator; ten had no formal leader.

Group II, Group condition: Self-contained-classroom teachers. Twenty groups of the same sizes as the teams described above were formed from the pool of self-contained-classroom teachers for the purpose of this experiment only. The seventy-five teachers making up these groups were from the same faculties but had never worked together as teaching teams. These groups were assumed to be ad hoc groups since the grouping was temporary and these teachers normally worked alone in a self-contained classroom. Nine of these groups were composed of primary-grade and eleven of intermediate-grade teachers.

TABLE 2

Characteristics of the Sample

Characteristics	Group Condition[a]		Individual Condition[b]		Total Sample	National Averages
	Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers		
Sex (%)						
Men	23	17	10	20	19	10
Women	77	83	90	80	81	90
Age (yrs.)						
Men	33.6	35.5	33.5	30.8	34.0	
Women	32.5	36.0	34.4	37.9	34.7	
Group	32.8	35.9	34.3	36.5	34.6	41.1
Teaching Experience (yrs.)						
Men	5.8	7.2	9.5	6.3	6.6	
Women	6.2	9.4	8.4	9.3	8.1	
Group	6.1	9.0	8.5	8.7	7.8	13.4
Masters Degrees (%)						
Men	41	38	50	25	39	
Women	2	8	6	31	8	
Group	11	13	10	30	14	15.7

[a] N=20 groups (75 individuals)

[b] N=20

Group III, Individual condition: Team teachers. Seven teams were randomly selected from the total pool of twenty-seven available teams and twenty teachers were randomly chosen from these teams to work in the individual problem-solving condition. Ten of these teachers were from three-member teams, five were from four-member teams, and five were from five-member teams. It was assumed that the teachers in Group III had had similar experiences, relative to team size, as team teachers placed in the group condition (Group I). Six of these teachers taught in primary-grade teams and fourteen in intermediate-grade teams.

Group IV, Individual condition: Self-contained-classroom teachers. Seven of the potential ad hoc groups were randomly selected from the total number of possible ad hoc groups and twenty teachers were randomly chosen from this pool to comprise the fourth experimental group. These teachers were from self-contained classrooms and had never had team-teaching experience. For the purpose of this experiment, these teachers worked as they normally did, in an individual decisicn-making situation. Eleven of these teachers taught a primary grade and nine taught an intermediate grade.

The Decision Process Test

Data were collected for all subjects* on two main classes of dependent variables: quality of decision, and decision-making behavior. These were measured by means of an instrument developed for this study and modeled on an instrument designed by Brim et al. (1962) for their study of parental decision making. The instrument consisted of two problems, each one describing a specific behavior of an elementary school boy. For each behavior, six alternative courses of teacher action were given. Subjects were asked to supply the possible consequences of each of these actions, to evaluate each one in terms of the foreseen consequences, and then to rank the six alternatives in order of their desirability (see Appendix A for sample pages of the instrument).

The Decision Process Test (DPT), like the instrument developed by Brim et al., was based on a normative decision model derived from formal decision theory (Taylor, 1965, pp. 48-86). This model assumes that a rational decision maker will consider all the known consequences of the alternatives open to him and will then choose the alternative offering him the maximum in utility or value. Consequences are evaluated in terms of the likelihood of their occurrence (probability) and their attractiveness to the decision maker (desirability). Brim et al. added another dimension to the evaluation of alternatives--a consideration of the time when each of the foreseen consequences might be expected to occur.

Although formal decision theory suggests that the decision maker will choose the alternative offering maximum personal utility, the subjects in this study were instructed to consider each teacher action in terms of the consequences to the child. Therefore, it was assumed in predicting outcomes that they would select among the alternatives so as to maximize the occurrence of consequences perceived as desirable for the child rather than for themselves. The decision model outlined above was not explained to them.

---

*Hereafter, "subject" refers either to a group working in a group condition or to an individual working in an individual condition.

Development of the DPT. A different sample of seventy-five elementary school teachers from both teams and self-contained classrooms identified problems frequently encountered with elementary school pupils on an open-ended questionnaire. They also ranked the problems that they had listed in order of perceived seriousness. From among the problems thus obtained, six were selected on the basis of frequency and the average rank given to seriousness. These problems were then submitted to another sample of seventy elementary school teachers who were asked to suggest a range (from "good" to "poor") of teacher actions in regard to each of the problems and also to evaluate each problem in terms of frequency of occurrence and seriousness.

Three problems and a range of six alternative teacher actions for each were selected for pretesting. In the pretesting situation, teachers were able to react to no more than two problems in an hour. Therefore, one problem, selected at random, was omitted from the final DPT. Wording was clarified on the final version of the problems and alternatives wherever pretesting had shown some confusion on the part of pretest subjects as to what was meant.

The DPT booklets were then made up for the study. Half of them began with one problem stated first; half with the other problem stated first. Within each of the problems, alternative teacher actions were arranged randomly so that all subjects did not see the alternatives in the same order. Standard directions for completing the DPT were prepared to be read to the subjects. The problems and the alternatives used in the DPT are outlined below.

Problem IB (Inattentive Behavior). Problem IB concerned the child's inattentive behavior and his difficulty with schoolwork because of his persistent failure to listen to explanations and directions for work. The alternative courses of teacher action given were as follows:

1. Seat the child near the front of the room so that he will be close to the teacher who is presenting a lesson or giving directions.
2. Whenever the child appears inattentive, have him repeat the teacher's question or the directions given.
3. Ignore the inattentive behavior.
4. Plan lessons for the entire class to teach the skills of listening and following directions.
5. Reward the child whenever he appears to be listening attentively or follows directions well.
6. Explain to the child why it is important for him to learn to listen and to follow directions.

Problem PA (Physical Aggressiveness). Problem PA involved the child's frequent displays of physical aggressiveness toward other children. The alternative teacher actions presented were:

1. Isolate the child from the children with whom he has trouble.
2. When he is not present, help the other children to understand his problem and let them suggest ways to help him.
3. Send him to the principal whenever this behavior occurs.
4. Let him express his aggressive feelings in more acceptable ways.
5. Exclude him from the class and/or playground whenever this behavior occurs.
6. Let the other children treat him in the same way so that he will find out how it feels.

Information base for decisions on the DPT. To give subjects a common basis for evaluating the given alternative teacher behaviors and the probable consequences of each, two folders of information about the student were prepared. One was for the fifth-grade level (intermediate); the other was for the second-grade level (primary). The folders were developed by using the second- and fifth-grade records of an actual child who frequently exhibited the given behaviors. All identifying information was changed. The California Cumulative Record folders for the elementary level were used. These folders contained report cards from previous grades; scores on achievement and mental maturity tests; information about parents and siblings; health records; a report of the school psychologist; record of absences; samples of schoolwork; and a photograph of the child.

Decision-behavior variables measured by the DPT. Twelve variables on the DPT measured the decision-making process. These included: (1) the total working time on each problem; (2) the number of consequences listed by subjects for alternative teacher actions; (3) the number of items of information from the cumulative folder used in reaching a decision; and (4) the number of additional courses of teacher action suggested. Another group of decision-process variables included: (1) the probability of perceived consequences as indicated by subjects; (2) the subjective desirability of perceived consequences; and (3) the probable time occurrence of perceived consequences. Each of these three variables was rated twice--once according to the direction and once according to the extremity of the subject's thinking (i.e., the frequency of use of the extreme ends of the scale for each variable). Table 3 shows the scoring system used.

TABLE 3

Scoring System for Variables of Probability, Desirability, and Time Occurrence of Consequences

Probability Direction		Desirability Direction		Time Direction		Extremity of Thinking[a]
Highly probable	.95	Strongly desire	2	Within a week	5	2
Probable	.75	Desire	1	A week to a month	4	1
Half and half	.50	Don't care either way	0	1 to 5 months	3	0
Improbable	.25	Do not desire	-1	6 months to a year	2	1
Very improbable	.05	Strongly do not desire	-2	One year or more	1	2

[a]Each variable was scored in the same manner.

The other two decision-process variables were measures of the rationality of the subjects' decision-making behavior. The investigator computed two expected utility scores for each alternative teacher action. The first was derived by multiplying the probability score by the desirability score and summing these for each alternative. The second expected utility score was computed in the same way except that time direction was also used as a factor in determining the expected utility. Alternative teacher behaviors were then ranked by the investigator according to each of these expected utility scores. The Kendall rank correlation coefficient, Tau, was computed for each of these rankings and the actual ranking given by the subject after evaluating each alternative. These correlations were taken as measures of the degree of rationality of decision behavior.

Decision-quality variables measured by the DPT. Five variables on the DPT were measures of the quality of the decision reached. The decision consisted of the final ranking given to the alternatives for each problem. The quality of this decision was determined by computing Kendall's rank correlation coefficient

between each subject's ranking and four different criterion rank orders. These were obtained by submitting the problems, the given alternatives, and the additional courses of teacher action suggested by subjects to a group of fifteen experts made up of five professors of elementary education, five educational psychologists, and five specialists in child growth and development. The rankings given by these experts were averaged as a total group and by subgroups to yield the four criterion rank orders (see Appendix B for a sample of the instrument submitted to the experts).

The experts also assessed the quality of the additional courses of action suggested by subjects after they had ranked the given alternatives. Each expert was asked to rate each suggested action by giving it a plus if he considered it superior to any of the given actions or a minus if he felt it was poorer than any of the given actions. Thus, if all fifteen experts rated a suggestion plus (+), the subject's score for that suggestion was 15; conversely, if all experts rated a suggestion minus (-), the score was -15. A composite score for all suggestions was computed for each subject.

The validity and reliability of the DPT. One of the objections to applying the findings of small-group research to established work groups is that most of this research has used contrived problems that do not generally occur in the actual daily tasks of such groups. In order to make the content of the test as task-related to teacher behavior as possible, the problems and alternative solutions used in the DPT were selected from among the problems and solutions actually suggested by elementary teachers. To determine whether the subjects actually saw the problems as relevant ones, they were asked to rate each problem according to frequency of occurrence and seriousness (see page 1 of the DPT, Appendix A). Responses were taken as a measure of the degree of subjective concern for the experimental task. In addition, they were asked to list the alternative teacher behaviors that they might actually perform, were they faced with similar real problems. From the answers to these questions, some indication of the face validity of the instrument was obtained.

No attempt was made to establish other types of test validity. Since no assessment was made of such factors as the intelligence of the subjects, teaching skill, personality, attitudes, and so forth, there was no way to determine which differences among subjects might account for observed differences and thus no way to evaluate the construct validity of the DPT. Nor was there a way to determine how subjects might handle the given kinds of problems in an actual situation; consequently predictive or concurrent validity could not be assessed. It was assumed, however, that the DPT provided optimum conditions for rational decision making--i.e., time, information, several alternatives, and an organized method for evaluating these components--and that individual differences

in test performance would correlate highly with individual differences among them in an actual decision situation.

The DPT yielded two scores for each subject on each of the seventeen variables--one for Problem IB and one for Problem PA. It was assumed that each score of a given pair was a measure of the same decision trait and that there would be some consistency in the way subjects responded to the same variable on each problem. However, no attempt was made to equate the problems and it was expected that the variance between the scores of each pair would be affected by the particular problem and its alternatives in addition to any unreliability of the test. Therefore, intercorrelations between scores were determined using the Pearson product-moment formula for the coefficient of correlation. The results were taken as estimates of the lower bounds of reliability on the various scores of the DPT.

The Hypotheses

Prior to collecting data, four main working hypotheses concerning probable outcomes were advanced. The first of these pertained to a main effect; the second to an interaction effect. The direction of the predicted difference on each was determined from prior small-group research and also from the decision model on which the DPT was based.

H_1: The decisions and the decision behavior of groups of teachers will be more like the criterion measures than will the decisions and decision behavior of individual teachers.

H_2: The decisions and the decision behavior of teams of teachers will be more like the criterion measures than will the decisions and decision behavior of ad hoc groups of teachers.

The criterion measure for superior decision behavior was defined as greater use of available information about the student from the cumulative folder; the listing of a greater number of consequences for the given courses of action; higher correlations between each of the two expected rank orders and the actual rank order of alternatives; and the listing of a greater number of additional courses of teacher action.* The criterion measure for the better decision was defined as a final rank order of alternatives correlating more closely with each of the four criterion rank orders; and higher-quality suggestions (as determined by the experts) for other suggested courses of teacher action.**

---

*Hereafter referred to as rational decision behavior.

**Hereafter referred to as decision quality.

A third hypothesis was concerned with a possible main effect and/or a possible interaction effect. There was no basis in prior educational research for predicting the direction of any such difference, however.

H_3: The decisions and decision behavior of team teachers (as groups or as individuals) will differ from the decisions and decision behavior of self-contained-classroom teachers.

Group I was composed of 20 experienced teams, 10 of which had appointed leaders and 10 of which did not. The fourth and final hypothesis concerned a possible difference between these two subgroups within this experimental condition. Although data are available on the effects of formal leadership within a small group, none were collected with teaching teams and little was known about the training, if any, that team leaders had received. Therefore, no direction was predicted for the possible difference:

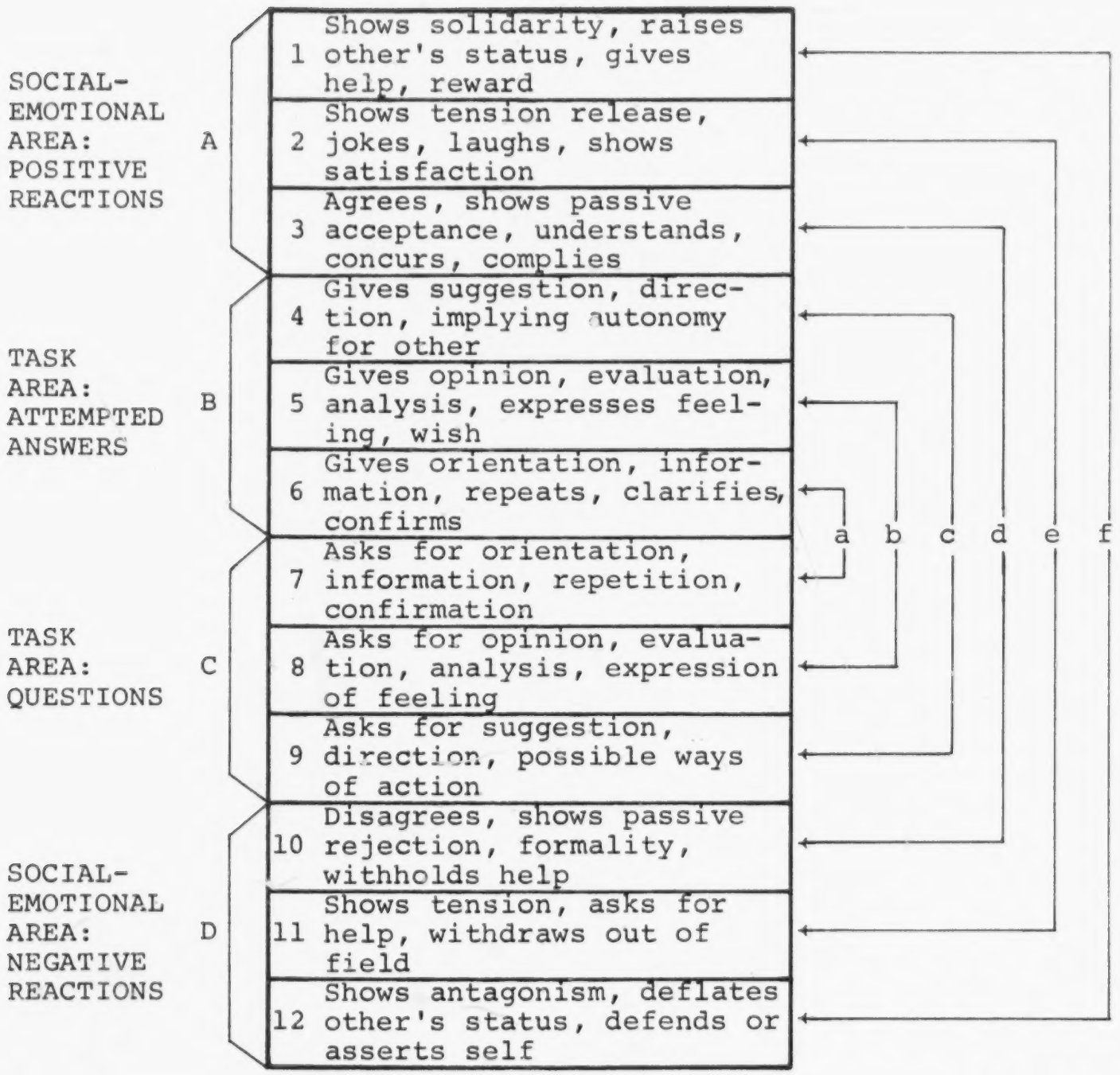
H_4: The decisions and the decision behavior of teams with appointed leaders will differ from the decisions and decision behavior of teams without appointed leaders.

Group Interaction

Small-group research provided a basis for hypothesizing that groups of teachers generally, and experienced teams especially, would handle the decision-making task better than teachers working individually. However, small-group research has identified many inhibiting effects of group work that may lower the actual quality of the group product. Therefore, a systematic observation of the groups' behavior was included in the design in order to aid in the interpretation of results of the DPT.

A system for categorizing group behavior. Data were collected on both the verbal and the apparent nonverbal behavior of subjects in each group condition. The system used for categorizing observed interaction was the Bales Interaction Process Analysis (see Figure 1). This system was chosen because it has clearly defined categories that have been used by Bales and numerous other investigators in studying the behavior of many different types of groups. The system also includes specific directions for training observers and for collecting data.

Training observers for this study. Audio tapes of actual team meetings, collected as part of the preliminary work for this study, were used to train the two observers. Verbal comments from typed transcriptions of these tapes were categorized by each trainee independently using Bales detailed descriptions of each category (Bales, 1950, pp. 177-95). Results were compared and differences resolved by consulting descriptions. When both


SOCIAL-EMOTIONAL AREA: POSITIVE REACTIONS
A
1 Shows solidarity, raises other's status, gives help, reward
2 Shows tension release, jokes, laughs, shows satisfaction
3 Agrees, shows passive acceptance, understands, concurs, complies
TASK AREA: ATTEMPTED ANSWERS
B
4 Gives suggestion, direction, implying autonomy for other
5 Gives opinion, evaluation, analysis, expresses feeling, wish
6 Gives orientation, information, repeats, clarifies, confirms
7 Asks for orientation, information, repetition, confirmation
TASK AREA: QUESTIONS
C
8 Asks for opinion, evaluation, analysis, expression of feeling
9 Asks for suggestion, direction, possible ways of action
SOCIAL-EMOTIONAL AREA: NEGATIVE REACTIONS
D
10 Disagrees, shows passive rejection, formality, withholds help
11 Shows tension, asks for help, withdraws out of field
12 Shows antagonism, deflates other's status, defends or asserts self
a b c d e f

Key: a, problems of orientation; b, problems of evaluation; c, problems of control; d, problems of decision; e, problems of tension-management; and f, problems of integration.

Fig. 1.--Bales' Interaction Process Categories. (Source: Robert F. Bales, Interaction Process Analysis [Cambridge, Mass.: Addison-Wesley, 1950], p. 625.)

trainees were interpreting the category system more or less in the same way, the categories were memorized by each and the trainees practiced with tape recordings to gain speed and accuracy.

To gain practice in noting nonverbal as well as verbal behavior and also in noting the originator of each behavior, the observers recorded the interaction of group members during six pretest trials with the DPT.

Collection of interaction data. As group members worked on the DPT, each observer attempted to record as much of the interaction of group members as possible. Large letters (A to E) with adhesive backs were worn by group members to enable observers to note the originator of an interaction. If several members of the group were talking at once (as in agreement with a statement of one member), the letter O was used to indicate that several people originated the given response. Nonverbal behavior was recorded simultaneously by each observer at one-minute intervals. A hidden timing device flashed a red light each minute to alert observers to notice and record nonverbal behavior. Tape recordings were made of each group session as a precaution against inadequate training of observers.

Methods of Data Analysis

One-way analysis of variance was used to test the null hypothesis of no difference among the four experimental groups on: the decision-behavior variables, the quality-of-decision variables, and the percentage of the groups' behavior in each of the behavior categories. Where the null hypothesis was rejected, the Tukey method for multiple comparisons was used to locate the contrasts responsible for the difference (see Guenther, 1964, pp. 54-57).

The main effects of group-versus-individual problem solving were determined by combining the means for experimental Groups I and II and comparing them with the combined means for experimental Groups III and IV. The main effects of the type of teaching assignment were determined by combining the means for experimental Groups I and III and comparing the results with the combined means for experimental Groups II and IV. Interaction effects were identified by comparing each sample mean with every other sample mean. A two-way analysis of variance for groups by size and type was also run on a few selected variables to determine whether the factor of group size was affecting results in any way. The following two sections contain the results of the above analyses.

THE DECISION PROCESS TEST: THE RESULTS OF DATA ANALYSIS

The results of the administration of the Decision Process Test were analyzed and summarized according to the four main working hypotheses.

Hypothesis 1. Groups Compared to Individuals

Decision behavior. In the first hypothesis it was postulated that groups would exhibit more rational decision behavior than individuals. Five variables served as indices of rational decision making:

1. the degree to which subjects' actual ranking of alternative teacher actions correlated with the expected ranking as determined by the investigator using subjects' scores on the variables of probability and desirability of perceived consequences of teacher actions (Table 4, variable 11)
2. the degree to which subjects' actual ranking correlated with the expected ranking based on subjects' scores on the variables of probability, desirability, and time occurrence of perceived consequences (Table 4, variable 12)
3. the use of more of the available information about the student
4. he listing of additional consequences for each alternative teacher action
5. the listing of additional possible courses of teacher action (Table 4, variables 2-4).

No support for the hypothesis was found on any of these variables. Differences were generally in the predicted direction for the first two of the above with both expected rankings tending to correlate more closely with actual rankings for groups than for individuals. However, mean correlations for all four experimental conditions tended to be rather low (range: 0.46-0.66). These data seem to indicate that no experimental group gave a great deal of consideration to the perceived consequences of teacher actions in terms of the probability of happening, desirability, and time occurrence of these in arriving at a final ranking. There were large variations within all experimental conditions, however, indicating that some subjects may have given more consideration than others to prior thinking in making the final ranking.

On the other variables assessing rational decision making (i.e., 3 to 5 above), results tended to favor the familiar mode of

TABLE 4

One-Way Analysis of Variance: Decision-Behavior Variables
(N=20)

Variable	Problem	Group Condition		Individual Condition		F-ratio
		Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers	
1. Decision time (in minutes)	IB Mean	40.65	33.05	28.85	29.05	3.05*
	SD	14.08	18.55	11.97	10.60	
	PA Mean	36.00	43.35	29.35	35.25	3.62*
	SD	13.94	17.75	11.31	9.53	
2. Total number of information items used in reaching decision (possible range: 0-60)	IB Mean	16.95	13.15	13.45	17.05	n.s.
	SD	11.00	11.64	9.89	13.58	
	PA Mean	14.90	12.90	13.30	15.40	n.s.
	SD	9.40	9.63	10.07	10.50	
3. Number of additional courses of action suggested	IB Mean	0.90	0.65	0.35	0.30	n.s.
	SD	1.29	0.93	0.57	0.74	
	PA Mean	1.15	0.80	0.35	0.95	n.s.
	SD	2.66	1.43	0.53	1.63	
4. Number of foreseen consequences (possible range: 6-30)	IB Mean	15.85	14.15	14.45	15.30	n.s.
	SD	4.42	6.00	4.12	6.19	
	PA Mean	16.30	15.15	15.20	17.50	n.s.
	SD	5.00	5.79	4.67	6.43	
5. Probability direction (possible range: .05-.95)	IB Mean	0.72	0.73	0.72	0.70	n.s.
	SD	0.07	0.06	0.11	0.11	
	PA Mean	0.77	0.72	0.75	0.70	3.06*
	SD	0.07	0.10	0.10	0.09	

6. Probability extremity (possible range: 0-2)	IB Mean	1.05		1.03		1.18		1.20		n.s.
	SD		0.29		0.23		0.45		0.29	
	PA Mean	1.26		1.08		1.17		1.14		n.s.
	SD		0.32		0.31		0.33		0.26	
7. Desirability direction (possible range: 2 to -2; negative scores indicate neg. reaction)	IB Mean	0.26		0.07		-0.11		0.09		n.s.
	SD		0.47		0.62		0.60		0.53	
	PA Mean	-0.40		-0.35		-0.43		-0.20		n.s.
	SD		0.53		0.46		0.53		0.46	
8. Desirability extremity (possible range: 0-2)	IB Mean	1.81		1.77		1.55		1.67		2.77*
	SD		0.24		0.25		0.27		0.44	
	PA Mean	1.86		1.83		1.50		1.66		8.48**
	SD		0.22		0.14		0.33		0.30	
9. Time direction (possible range: 1-5; scores indicate present time PA orientation)	IB Mean	4.40		4.44		4.39		4.08		n.s.
	SD		0.35		0.35		0.44		1.03	
	PA Mean	4.43		4.48		4.26		4.30		n.s.
	SD		0.40		0.25		0.50		0.34	
10. Time extremity (possible range: 0-2)	IB Mean	1.41		1.47		1.44		1.37		n.s.
	SD		0.35		0.33		0.43		0.43	
	PA Mean	1.49		1.48		1.34		1.37		n.s.
	SD		0.36		0.25		0.40		0.32	
11. Correlation of actual rank ordering with utility rank ordering (P x D)[a]	IB Mean	0.56		0.57		0.49		0.46		n.s.
	SD		0.30		0.25		0.36		0.27	
	PA Mean	0.66		0.47		0.52		0.49		n.s.
	SD		0.17		0.27		0.30		0.29	
12. Correlation of actual rank ordering with utility rank ordering (P x D x T)[a]	IB Mean	0.54		0.57		0.47		0.46		n.s.
	SD		0.32		0.23		0.34		0.27	
	PA Mean	0.64		0.47		0.48		0.49		n.s.
	SD		0.22		0.24		0.31		0.32	

*p <.05. **p < .01.

[a]Using Kendall's rank order correlation coefficient.

work rather than to be in the direction of the hypothesis. Thus teams working as groups and self-contained-classroom teachers working as individuals tended to use more of the available information, list more consequences for each alternative and more other courses of action than did ad hoc groups or team teachers working alone.

In addition to the above comparisons, groups were also contrasted with individuals on the six variables from which the scores used by the investigator in computing expected rankings were obtained. Again, these were: direction (high or low) of the probability of perceived consequences happening; direction (positive or negative) of the desirability of perceived consequences; direction (present or future) of time occurrence of perceived consequences, and the rating of extremity on probability; desirability; and time (Table 4, variables 5-10). One other comparison was made on the total decision time required by subjects (Table 4, variable 1). One significant difference between groups and individuals and two significant differences between one experimental condition and another were found on these variables.

The significant difference between groups and individuals was on the variable of desirability extremity (or the frequency of use of the extreme ends of the desirability scale) for Problem PA (.01). Groups were more extreme in the way they rated the desirability of the perceived consequences of teacher actions they had listed. Although groups offered essentially the same consequences for the given teacher actions as did individuals, they saw these consequences as being better or worse for the child. On Problem IB, teaching teams were also more extreme in rating the desirability of perceived consequences of alternatives than were team teachers who worked in the individual conditions (.05), which tends to support the finding on Problem PA. Generally, all subjects were rather extreme on this variable. A range from 0 to 2.00 was possible with the overall mean score for groups being 1.82 and for individuals 1.59 and with standard deviations of 0.21 and 0.33, respectively. These subjects seemed to be mainly concerned with consequences for the child which were either very good or very poor in their opinions.

Teaching teams also regarded the consequences they had listed for Problem PA as more likely to happen than did individuals from self-contained classrooms. Although this was the only significant difference noted on this variable, the common tendency of all subjects was to list consequences that they saw as quite likely to happen. The mean probability direction across all experimental conditions on both problems was 0.73, with a mean standard deviation of 0.09. Very little consideration appeared to be given, generally, to the less probable consequences for the child of the given teacher actions.

Another common tendency on the part of all subjects, regardless of experimental condition, was to list consequences that they saw as likely to occur within the immediate future (Table 4, variable 9). The possible range was from 1 to 5, with a score of 1 representing consequences not likely to occur for a year or more and a score of 5 representing those expected within a week. The overall mean score on both problems was 4.35, with an average standard deviation of 0.46. Apparently, remote consequences to the child because of the given teacher actions did not occur to subjects.

Although there were no significant differences between groups and individuals on total decision time, groups generally took longer to work on the problems than did individuals (Table 4, variable 1). A suggested working time of one hour was given all subjects. Individuals tended to stay within this limit, while groups needed about fifteen minutes longer to complete the test. There were also two significant differences between experimental conditions with teams working longer on Problem IB than did individuals from teams (.05) and ad hoc groups working longer on Problem PA than did individuals from teams (.05).

Decision quality. No significant differences were found to support the hypothesis that groups would make better decisions than individuals. However, differences were in the predicted direction and two significant effects were found between experimental conditions. On Problem IB, ad hoc groups differed from individuals from self-contained classrooms (.05) in the degree of correlation with educational psychologists. On Problem PA, teams differed from individuals from teaching teams (.01) in the degree of correlation with professors of elementary education (Table 5, variables 3-4). Correlations with educational psychologists tended to be higher than with other experts (see Appendix B, Table B-2 for degree of agreement among experts on rank ordering of alternatives for each problem).

There were no significant differences when groups were compared to individuals on the scores for quality of additional courses of action suggested (Table 5, variable 5). However, group scores were higher than individual scores on both problems with differences averaging about 3 score points. On Problem IB, the F test rejected the null hypothesis of no differences among experimental groups. However, no significant contrasts were located using Tukey's method for multiple comparisons. A difference of 4.00 in group means would have been necessary to reach the .05 level of significance. The difference in means between teams as groups and individuals from teams was 3.75, while the difference between ad hoc groups and individuals from teams also approached the necessary amount at 3.65.

TABLE 5

One-Way Analysis of Variance: Decision-Quality Variables
(N=20)

Variable	Problem	Group Condition		Individual Condition		F-ratio
		Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers	
1. Correlation of subject's rank ordering with average rank ordering of total group of experts[a]	IB Mean	0.60	0.62	0.52	0.48	n.s.
	SD	0.25	0.12	0.23	0.32	
	PA Mean	0.65	0.61	0.50	0.61	n.s.
	SD	0.16	0.16	0.27	0.21	
2. Correlation of rank ordering with average rank ordering by child growth and development specialists[a]	IB Mean	0.54	0.59	0.51	0.46	n.s.
	SD	0.30	0.19	0.25	0.29	
	PA Mean	0.49	0.47	0.40	0.51	n.s.
	SD	0.16	0.16	0.24	0.26	
3. Correlation of subject's rank ordering with average rank ordering by educational psychologists[a]	IB Mean	0.49	0.65	0.54	0.43	3.41*
	SD	0.27	0.16	0.24	0.23	
	PA Mean	0.70	0.72	0.61	0.62	n.s.
	SD	0.21	0.20	0.34	0.16	
4. Correlation of subject's rank ordering with average rank ordering by professors of elementary education[a]	IB Mean	0.51	0.61	0.44	0.43	n.s.
	SD	0.26	0.19	0.31	0.30	
	PA Mean	0.57	0.43	0.31	0.48	5.90**
	SD	0.18	0.15	0.24	0.20	
5. Scores on quality of additional courses of action suggested by subject	IB Mean	4.05	3.95	0.30	1.50	2.98*
	SD	5.73	6.29	2.72	3.59	
	PA Mean	8.05	5.15	2.95	6.15	n.s.
	SD	11.53	8.50	7.29	10.79	

*$p < .05$. **$p < .01$.

[a]Using Kendall's rank order correlation coefficient.

Hypothesis 2. Teams Compared to Ad Hoc Groups

No significant differences on decision-behavior or decision-quality variables were found to support the hypothesis that teams would do better than ad hoc groups of self-contained-classroom teachers when these two experimental groups were compared by one-way analysis of variance. Since the group condition contained groups of three different sizes, a two-way analysis of variance for groups by size and type was run on certain variables to determine whether the size of a group might be masking differences that were due to the type of group. The decision-behavior variables selected for this analysis were: desirability direction; desirability extremity; and correlation between expected and actual rankings (P x D and P x D x T). The decision-quality variable was the correlation of the average ranking of alternatives by the total group of experts with the ranking by subjects.

Groups by size and type: Decision behavior. On the variable of desirability direction, there was a significant interaction effect (.05) between size and type of group on Problem IB and also a trend toward significance (.10) for the type of group to make a difference (see Tables 6 and 7). Teams became more positive in evaluating the desirability of perceived consequences as team size increased; ad hoc groups became more negative.

On the variable of desirability extremity, there was a trend for the size of group to make a difference ($p < .25$) with both types of groups becoming more extreme as the size of the group increased. Although this was an insignificant trend, it is noted because there was a main effect between groups and individuals on this variable.

On Problem PA, there was a significant difference (.05) between teams and ad hoc groups on the degree of correlation between subjects' actual ranking and the expected ranking of alternative teacher actions as computed by the investigator. Teams displayed more rational behavior, i.e., tended to rank order alternatives more as they were expected to rank them given their scores on the variables of probability, desirability, and time occurrence of perceived consequences. Apparently, group size did tend to mask this difference in the one-way analysis of variance on this variable.

Groups by size and type: Decision quality. One comparison was made between teams and ad hoc groups on the quality of their decisions. This was for Problem PA on the degree of correlation between subjects' ranking of alternatives and experts' average ranking. There was only a trend ($p < .10$) toward an interaction effect between size and type of group with teams tending to have lower correlations as the size of the group increased while ad hoc groups tended to have higher correlations as the size of the group increased.

TABLE 6

Two-Way Analysis of Variance for Groups by Size and Type: Decision-Behavior Variables

Variable	Source	Problem	SS	df	F
Desirability direction	Group size	IB	0.42	2	0.78
	Group type	IB	0.97	1	3.65‡
	Interaction	IB	2.05	2	3.83*
	Error	IB	9.07	34	
	Total	IB	12.51	39	
Desirability extremity	Group size	IB	0.23	2	2.02‡
		PA	0.08	2	1.68‡
	Group type	IB	0.00	1	0.01
		PA	0.00	1	0.12
	Interaction	IB	0.12	2	1.12
		PA	0.00	2	0.08
	Error	IB	1.96	34	
		PA	1.17	34	
	Total	IB	2.31	39	
		PA	1.25	39	
Correlation of actual and expected ranking (P x D)	Group size	PA	0.18	2	1.88‡
	Group type	PA	0.29	1	5.92*
	Interaction	PA	0.06	2	0.69
	Error	PA	1.64	34	
	Total	PA	2.17	39	
Correlation of actual and expected ranking (P x D x T)	Group size	PA	0.08	2	0.73
	Group type	PA	0.28	1	5.13*
	Interaction	PA	0.05	2	0.50
	Error	PA	1.86	34	--
	Total	PA	2.27	39	

*$p < .05$.

‡$p < .25$.

TABLE 7

Means and Standard Deviations for Groups by Size and Type: Decision-Behavior Variables

Variable	Members in Group	Problem	Teaching Teams		Ad hoc Groups	
Desirability direction	3	IB Mean	0.15		0.38	
		SD		0.55		0.60
	4	IB Mean	0.28		-0.08	
		SD		0.33		0.60
	5	IB Mean	0.46		-0.39	
		SD		0.43		0.35
Desirability extremity	3	IB Mean	1.80		1.67	
		SD		0.18		0.31
		PA Mean	1.84		1.74	
		SD		0.25		0.13
	4	IB Mean	1.71		1.85	
		SD		0.38		0.15
		PA Mean	1.83		1.82	
		SD		0.27		0.17
	5	IB Mean	1.94		1.90	
		SD		0.14		0.10
		PA Mean	1.94		1.94	
		SD		0.10		0.05
Correlation of actual and expected ranking (P x D)	3	PA Mean	0.69		0.49	
		SD		0.17		0.25
	4	PA Mean	0.65		0.60	
		SD		0.17		0.16
	5	PA Mean	0.59		0.31	
		SD		0.16		0.35
Correlation of actual and expected ranking (P x D x T)	3	PA Mean	0.62		0.48	
		SD		0.24		0.25
	4	PA Mean	0.67		0.57	
		SD		0.23		0.17
	5	PA Mean	0.64		0.35	
		SD		0.21		0.24

Hypothesis 3. Team Teachers Compared to Self-Contained-Classroom Teachers

The data did not support the supposition that team teachers would differ from self-contained-classroom teachers either in decision making or in the quality of the decision reached. In a total of thirty-four comparisons, seventeen on each problem, one difference was found between individual teachers from teaching teams and individuals from self-contained classrooms. When the correlations of each of these groups with professors of elementary education were compared by one-way analysis of variance, they were found to differ significantly (.05). However, considering that there was only one such difference, this probably should be attributed to random variation.

Hypothesis 4. Teams with Leaders Compared to Teams without Leaders

The final stated purpose of this study was to look at differences among teams that were related to the leadership factor. Therefore, the ten teams with appointed leaders were compared to the ten teams without appointed leaders on decision-making behavior and quality of decision. No significant differences were found on either class of variables although three differences on Problem PA approached significance (.10). These are reported since this exploratory part of the study was concerned with locating some promising leads for further research on team teaching. The following tendencies were discovered (see Table 8):

1. Teams without leaders tended to think of more consequences for alternative teacher behaviors.
2. Teams without leaders tended to be more rational.
3. The decisions of teams without leaders tended to correlate better with those of the experts.

The Validity and Reliability of the Decision Process Test

The degree to which the above findings represent meaningful comparisons among experimental groups depends in part on the validity and reliability of the DPT as a measure of the decision process. Therefore, an attempt was made to assess both of these factors.

Validity. The content of the DPT consisted of the problems themselves and the alternative teacher actions given for each problem. These were selected from a number of such problems and alternatives suggested by two independent samples of elementary school teachers. In order to assess the content validity of the DPT, however, it was necessary to determine whether the subjects in this study saw these problems as meaningful ones and at least some of

TABLE 8

One-Way Analysis of Variance: Decision-Behavior Variables
(N = 10)

Variable	Problem	Teams with Leaders		Teams without Leaders		F-ratio
Number of foreseen consequences	IB Mean	14.70		17.00		
	SD		5.08		3.53	n.s.
	PA Mean	14.20		18.40		
	SD		5.67		3.27	4.11*
Correlation of actual and expected ranking (P x D)	IB Mean	0.58		0.54		
	SD		0.28		0.33	n.s.
	PA Mean	0.59		0.72		
	SD		0.16		0.16	3.38*
Correlation of actual and expected ranking (P x D x T)	IB Mean	0.58		0.62		
	SD		0.20		0.29	n.s.
	PA Mean	0.58		0.70		
	SD		0.19		0.10	2.97*

*p < .10.

the alternative teacher behaviors as acceptable actions to take given the type of problems presented.

The content validity of the problems was determined by asking subjects to indicate their familiarity and concern with the given problems. Appropriate responses were checked along two five-point scales, one for occurrence of the given student behavior in subjects' experience and one for their rating of the seriousness of such behavior (see Appendix A for rating scales). The experimental conditions did not differ significantly in the way subjects perceived these problems. Table 9 indicates that all subjects appeared to be familiar with such student behavior and more than moderately concerned about it.

In order to determine if the alternative teacher actions were realistic ones for these subjects, they also were asked to indicate which of the actions they would take if actually faced with similar student behavior. In developing the DPT an effort was made to obtain a range of teacher actions, i.e., some "good" ways of handling the problem and some "poor" ways. Three of the alternatives given for each problem had been rated good and three had been rated poor by the sample of teachers from which these were

TABLE 9

Subjects' Ratings of Occurrence and Seriousness of Problems

Group Mean	Problem IB		Problem PA	
	Occurrence	Seriousness	Occurrence	Seriousness
Teams	2.8	2.6	3.1	2.3
Ad hoc Groups	2.4	2.2	3.2	2.0
Individuals from Teams	2.5	2.3	2.9	2.1
Individuals from self-contained classrooms	2.8	2.3	3.3	2.2
Mean	2.6	2.4	3.1	2.2

Note: 3.0 = "some" occurrences and "moderate" seriousness. Low scores indicate greater seriousness or higher frequency.

drawn. Table 9 shows the number of times each alternative was listed by subjects in each experimental condition of this study as an acceptable action for teachers to take. The average number of alternatives selected out of the six given is also shown in Table 10. It appears from this table that subjects in this study judged the quality of these teacher actions in approximately the same way as the teachers who suggested the alternatives.

Tables 9 and 10 seem to indicate that the DPT presented the subjects in this study with problems which were perceived as valid and a range of alternatives about each problem, some of which approached the level of the probable actions they would take if faced with such problems.

Reliability. Lower-bound estimates of reliability on the DPT were obtained by computing the Pearson product-moment coefficient of correlation between the score on each variable of Problem IB and the corresponding score on the same variable for Problem PA. Table 10 shows the inter-problem correlations on decision-process variables; Table 11 on decision-quality variables. As these tables indicate, individuals apparently were more consistent from problem to problem than were groups. Scores on the two problems were

TABLE 10

Teacher Behaviors Subjects Would Actually Enact
(number of times listed)

Alternatives	Group Condition		Individual Condition	
	Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers
Problem IB				
Seat the child near the front of the room	17	16	14	16
Have the child repeat the teacher's question	6	7	7	10
Ignore the inattentive behavior	5	2	3	1
Plan lessons for the entire class or group to teach skills of listening	11	19	9	11
Reward the child whenever he appears to be listening	17	20	17	18
Explain to the child why it is important to listen	16	15	13	17
Average number of alternatives listed as acceptable	3.6	4.0	3.2	3.7
Problem PA				
Isolate the child	16	17	11	11
Help the other children to understand	14	15	12	12
Send him to the principal	5	6	8	4
Let him express aggression in more acceptable ways	19	19	16	15
Exclude him from class and/or playground	9	9	10	8
Let others treat him the same way	1	1	3	1
Average number of alternatives listed as acceptable	3.2	3.4	3.0	2.6

TABLE 11

Decision Behavior: Correlation Between DPT Scores on Problems IB and PA

Variable	Group Condition		Individual Condition	
	Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers
Decision time	-.89*	-.87*	-.52*	-.82*
Total number of information items used in reaching decision	.79*	.62*	.84*	.98*
Number of foreseen consequences	.22	-.06	.48*	.68*
Number of additional courses of action suggested	.82*	.31	.38*	.30
Probability direction	.32	.25	.67*	.67*
Probability extremity	.38*	.11	.66*	.42*
Desirability direction	.55*	.05	.44*	.14
Desirability extremity	.74*	.28	.80*	.62*
Time direction	.24	.25	.56*	.92*
Time extremity	.20	.22	.59*	.32
Correlation of actual with expected ranking (P x D)	.76*	-.24	-.34	.39*
Correlation of actual with expected ranking (P x D x T)	.42*	-.10	-.40*	.44*

*$p < .05$.

significantly correlated for twelve out of the seventeen variables for each of the two individual conditions. Only two significant correlations were found for ad hoc groups, while eight were found for teams. For both groups and individuals, correlations were significant more often on decision-process variables than on decision-quality variables.

Although Tables 11 and 12 seem to indicate that the DPT was not very reliable for groups, it was probably more so than the

lower-bound estimates shown in these tables would suggest. Low correlations can be due to factors other than low test reliability. One such factor is a differential reaction to the two problems for which scores were correlated. One variable which provides some indication of such a differential reaction is that of desirability direction, or the degree to which subjects felt they liked the consequences of any given teacher alternative. Subjects in both the group condition and the individual condition were more negative toward the alternative teacher behaviors given for Problem PA than for Problem IB, but the range between problems on this variable was greater for groups than for individuals. Therefore, it seems reasonable to assume that the variance between scores on the same variable for the two problems was due, at least in part, to an interaction between the subject and the particular problem and not totally to low test reliability.

TABLE 12

Decision Quality: Correlation Between DPT Scores on Problems IB and PA

Variable	Group Condition		Individual Condition	
	Teams	Ad hoc Groups	Team Teachers	S-C-C Teachers
Correlation of subjects' rank ordering with average rank ordering of total group of experts	-.13	-.03	-.34	.75*
Correlation of rank ordering with average rank ordering by child growth and development specialists	-.19	.12	.44*	.26
Correlation of rank ordering with average rank ordering by educational psychologists	.10	-.03	-.21	.41*
Correlation of rank ordering with average rank ordering by professors of elementary education	.13	.29	-.15	.07
Scores on quality of additional courses of action suggested by subject	.21	.17	.32	.61*

*$p < .05$.

GROUP INTERACTION: THE RESULTS OF DATA ANALYSIS

The interaction of group members in the two group conditions (teams and ad hoc groups) was observed and their behavior analyzed using Bales Interaction Analysis for the purpose of further interpretation of results on the Decision Process Test. The results of the various analyses will be presented before the relationship of these to results on the Decision Process Test is shown.

Distribution of Group Behavior: Pooled Interaction Profiles

There were ten three-member groups, five four-member groups, and five five-member groups among the teams and also among the ad hoc groups. Disregarding group size, a profile was made of the typical, or average, distribution of behavior over the twelve Bales categories in the twenty teams (Figure 2) and also in the twenty ad hoc groups (Figure 3). Disregarding the type of group--i.e., whether team or ad hoc group--groups were divided into (five-member) and small (three-member), and profiles were made of the typical distribution of behavior (Figures 4 and 5). Profiles were also made to compare teams with leaders to teams without leaders (see Figures 6 and 7).

Task-related behavior. Regardless of whether teams were compared to ad hoc groups, or large groups to small ones, or teams with leaders to teams without leaders, approximately 74 percent of the total behavior of group members was related to the task with a range among groups from 69 to 81 percent. Within the task-related categories, Category 5 (Gives opinion) was used most frequently, followed by Category 6 (Gives orientation); the others (Categories 4, 7-9) were seldom used. Seventy-four percent of the behavior was concerned with the task, and of this amount approximately 60 percent consisted of attempted answers. Apparently, most of these answers did not come in response to questions raised by other group members since the question categories (7-9) only accounted for approximately 14 percent of the task behavior. However, in this study, large groups tended to ask more questions than small groups, probably because it was more difficult for each group member to see the information in the cumulative folder.

Social behavior. Approximately 26 percent of the behavior of groups, regardless of size or type, was classified as socio-emotional, either positive or negative, with a range among groups from 19 to 30 percent. Positive socio-emotional reactions were more frequent than negative ones with a range from 17 to 22 percent; negative reactions ranged from 5 to 8 percent. The most frequently used category was Category 3 (Agrees), followed by Category 1 (Shows solidarity). The most frequently used category of negative socio-emotional behavior was Category 10 (Disagrees). Negative socio-emotional behavior tended to occur more often in

1. Shows solidarity, raises other's status, gives help, reward	
2. Shows tension release, jokes, laughs, shows satisfaction	
3. Agrees, shows passive acceptance, understands, concurs, complies	
4. Gives suggestion, direction, implying autonomy for other	
5. Gives opinion, evaluation, analysis, expresses feeling, wish	
6. Gives orientation, information, repeats, clarifies, confirms	
7. Asks for orientation, information, repetition, confirmation	
8. Asks for opinion, evaluation, analysis, expression of feeling	
9. Asks for suggestion, direction, possible ways of action	
10. Disagrees, shows passive rejection, formality, withholds help	
11. Shows tension, asks for help, withdraws out of field	
12. Shows antagonism, deflates other's status, defends or asserts self	

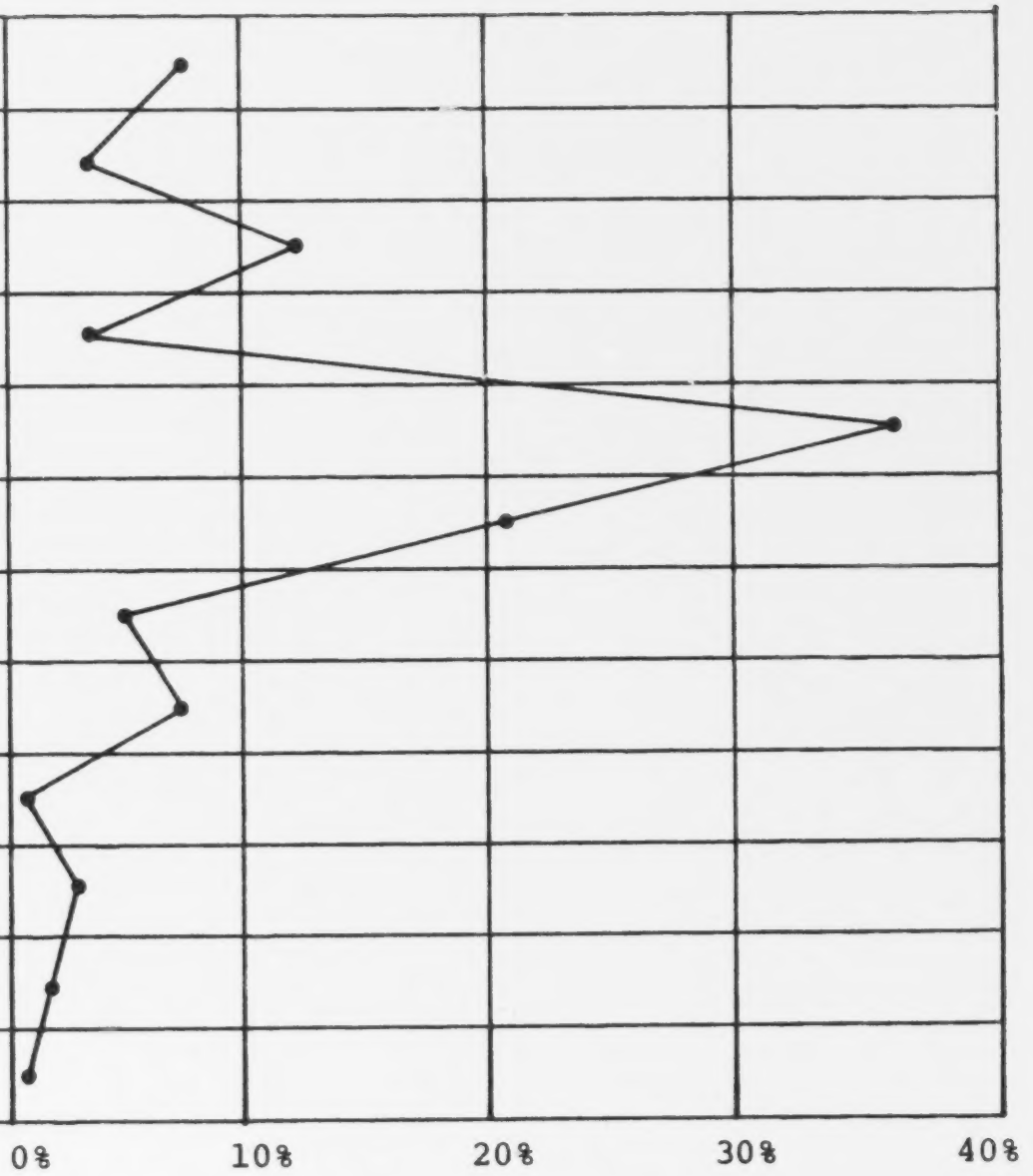


Fig. 2.--Pooled Interaction Profile: Teams (N=20)

1. Shows solidarity, raises other's status, gives help, reward	
2. Shows tension release, jokes, laughs, shows satisfaction	
3. Agrees, shows passive acceptance, understands, concurs, complies	
4. Gives suggestion, direction, implying autonomy for others	
5. Gives opinion, evaluation, analysis, expresses feeling, wish	
6. Gives orientation, information, repeats, clarifies, confirms	
7. Asks for orientation, information, repetition, confirmation	
8. Asks for opinion, evaluation, analysis, expression of feeling	
9. Asks for suggestion, direction, possible ways of action	
10. Disagrees, shows passive rejection, formality, withholds help	
11. Shows tension, asks for help, withdraws out of field	
12. Shows antagonism, deflates other's status, defends or asserts self	

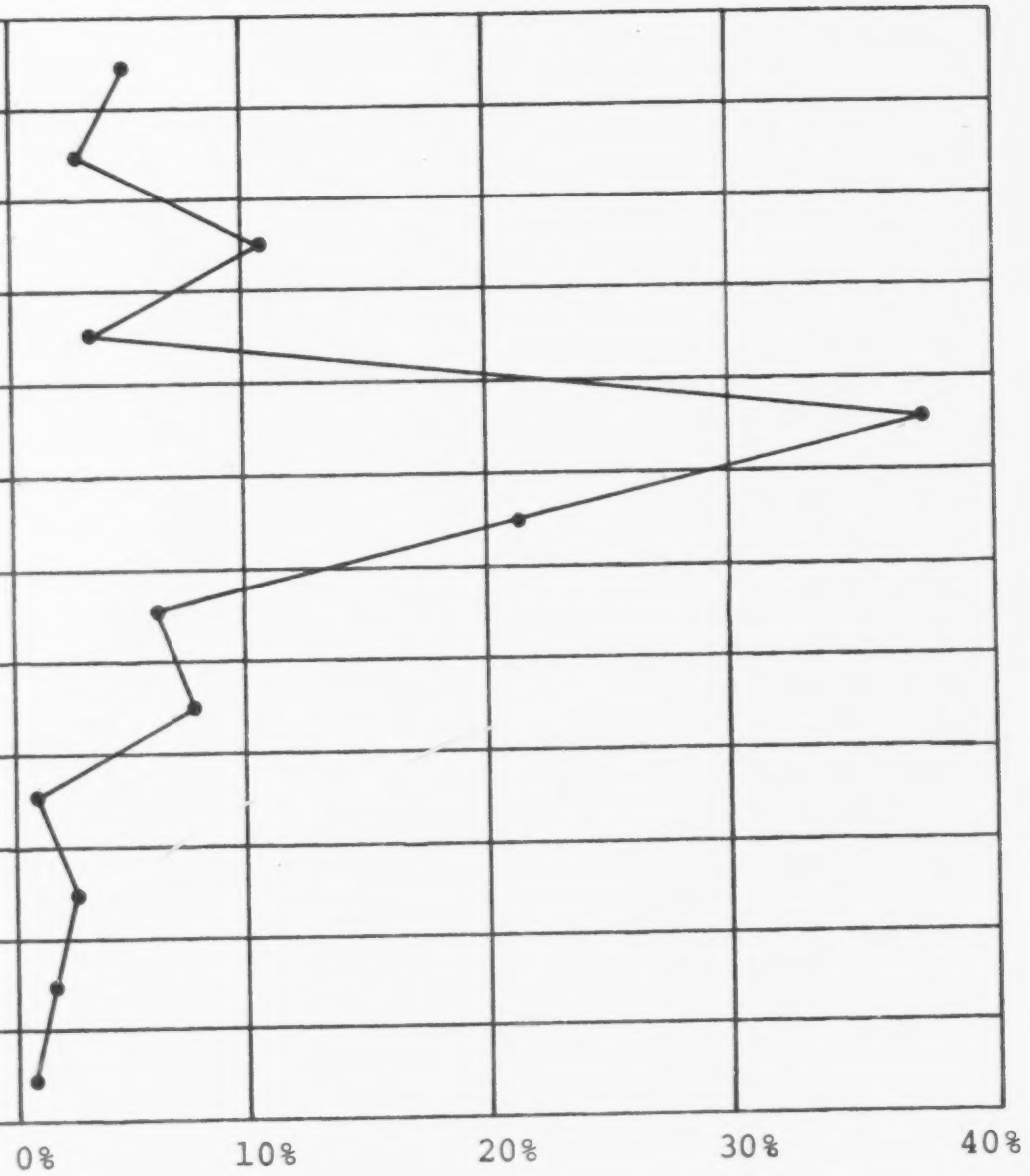


Fig. 3.--Pooled Interaction Profile: Ad Hoc Groups (N=20)

1.	Shows solidarity, raises other's status, gives help, reward
2.	Shows tension release, jokes, laughs, shows satisfaction
3.	Agrees, shows passive acceptance, understands, concurs, complies
4.	Gives suggestion, direction, implying autonomy for others
5.	Gives opinion, evaluation, analysis, expresses feeling, wish
6.	Gives orientation, information, repeats, clarifies, confirms
7.	Asks for orientation, information, repetition, confirmation
8.	Asks for opinion, evaluation, analysis, expression of feeling
9.	Asks for suggestion, direction, possible ways of action
10.	Disagrees, shows passive rejection, formality, withholds help
11.	Shows tension, asks for help, withdraws out of field
12.	Shows antagonism, deflates other's status, defends or asserts self

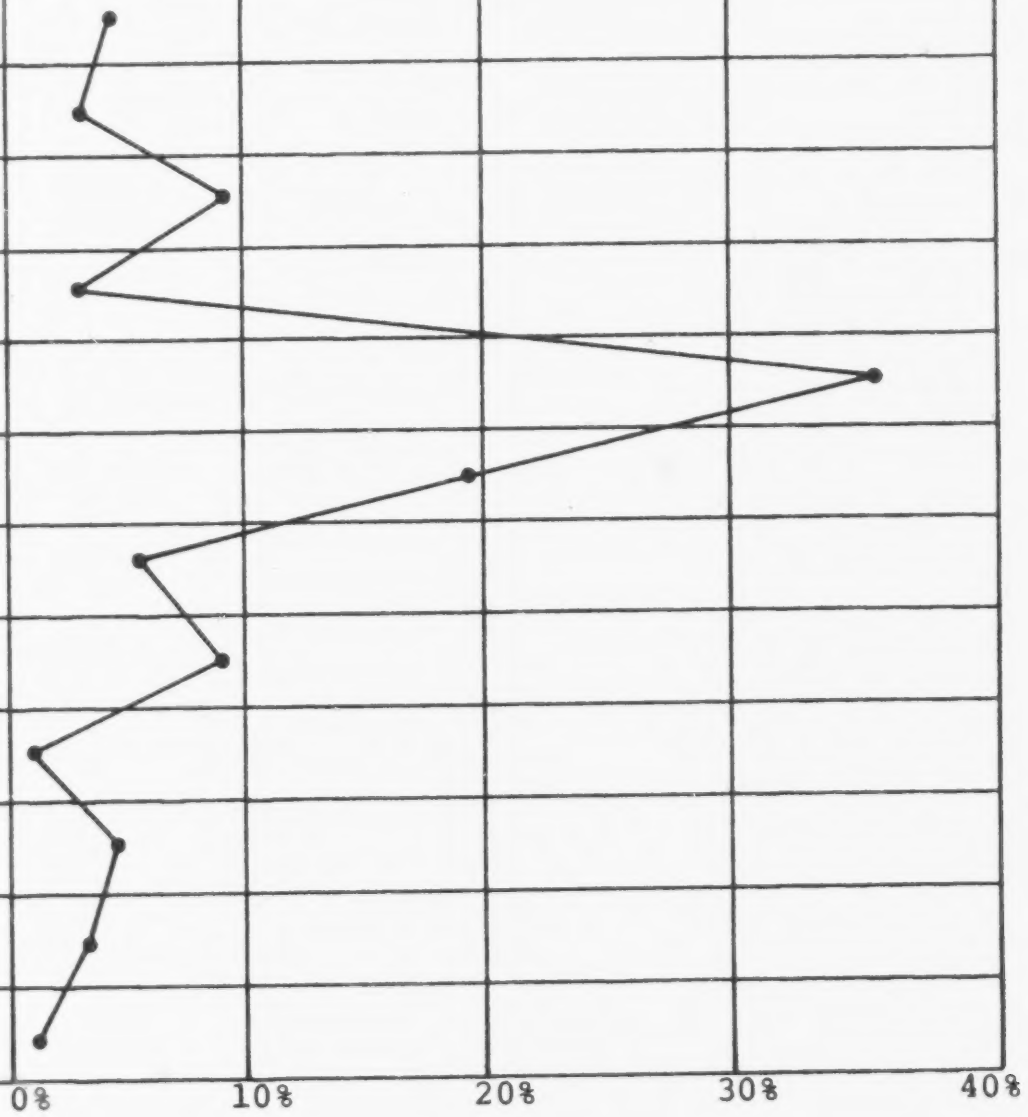


Fig. 4.--Pooled Interaction Profile: Five-member Groups (N=10)

1. Shows solidarity, raises other's status, gives help, reward	
2. Shows tension release, jokes, laughs, shows satisfaction	
3. Agrees, shows passive acceptance, understands, concurs, complies	
4. Gives suggestion, direction, implying autonomy for others	
5. Gives opinion, evaluation, analysis, expresses feeling, wish	
6. Gives orientation, information, repeats, clarifies, confirms	
7. Asks for orientation, information, repetition, confirmation	
8. Asks for opinion, evaluation, analysis, expression of feeling	
9. Asks for suggestion, direction, possible ways of action	
10. Disagrees, shows passive rejection, shows formality, withholds help	
11. Shows tension, asks for help, withdraws out of field	
12. Shows antagonism, deflates other's status, defends or asserts self	

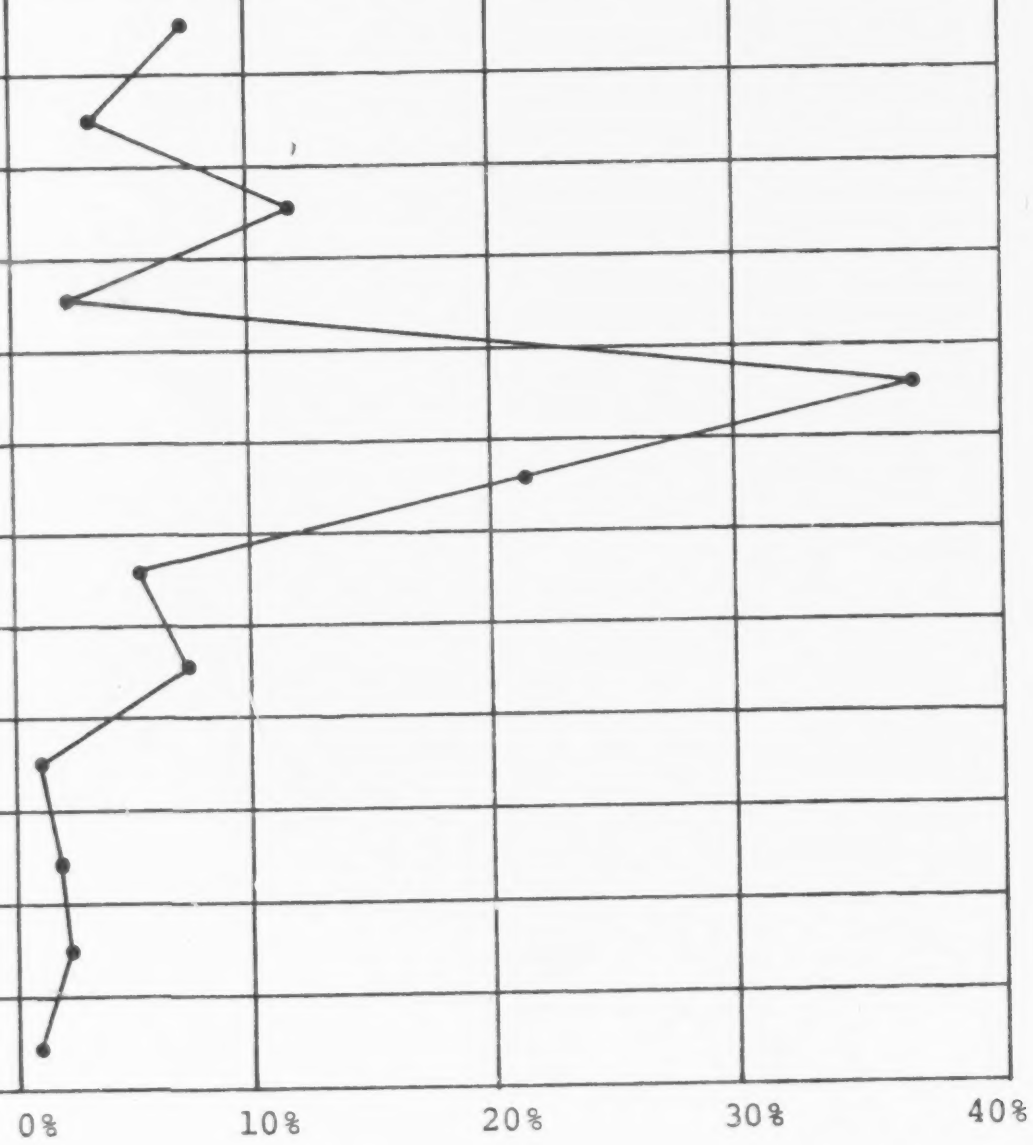


Fig. 5.--Pooled Interaction Profile: Three-member Groups (N=20)

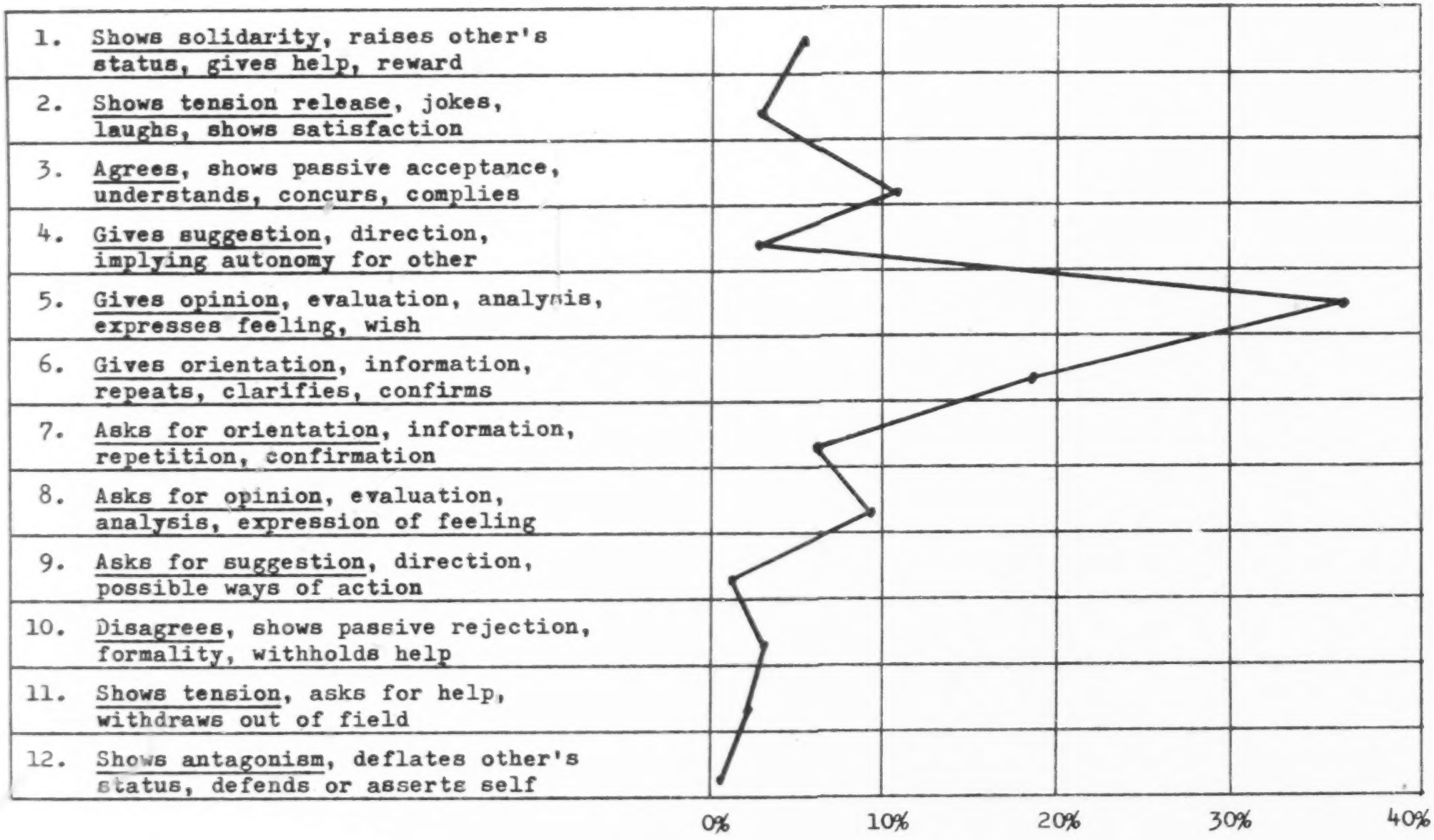

1. Shows solidarity, raises other's status, gives help, reward
2. Shows tension release, jokes, laughs, shows satisfaction
3. Agrees, shows passive acceptance, understands, concurs, complies
4. Gives suggestion, direction, implying autonomy for other
5. Gives opinion, evaluation, analysis, expresses feeling, wish
6. Gives orientation, information, repeats, clarifies, confirms
7. Asks for orientation, information, repetition, confirmation
8. Asks for opinion, evaluation, analysis, expression of feeling
9. Asks for suggestion, direction, possible ways of action
10. Disagrees, shows passive rejection, formality, withholds help
11. Shows tension, asks for help, withdraws out of field
12. Shows antagonism, deflates other's status, defends or asserts self
0%
10%
20%
30%
40%

Fig. 6.--Pooled Interaction Profile: Teams with Leaders (N=10)

1. Shows solidarity, raises other's status, gives help, reward	
2. Shows tension release, jokes, laughs, shows satisfaction	
3. Agrees, shows passive acceptance, understands, concurs, complies	
4. Gives suggestion, direction, implying autonomy for other	
5. Gives opinion, evaluation, analysis, expresses feeling, wish	
6. Gives orientation, information, repeats, clarifies, confirms	
7. Asks for orientation, information, repetition, confirmation	
8. Asks for opinion, evaluation, analysis, expression of feeling	
9. Asks for suggestion, direction, possible ways of action	
10. Disagrees, shows passive rejection, formality, withholds help	
11. Shows tension, asks for help, withdraws out of field	
12. Shows antagonism, deflates other's status, defends or asserts self	

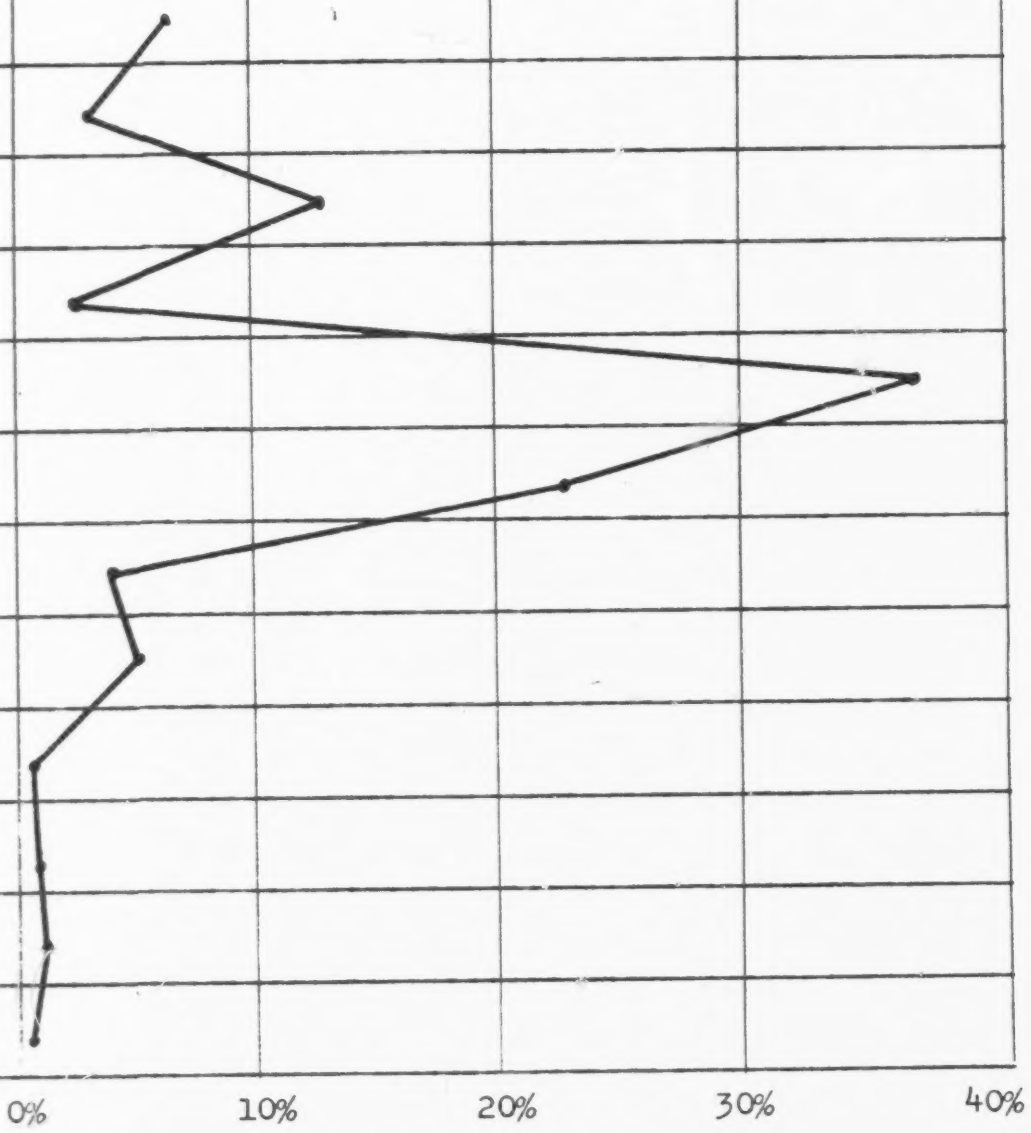


Fig. 7.--Pooled Interaction Profile: Teams without Leaders (N=10)

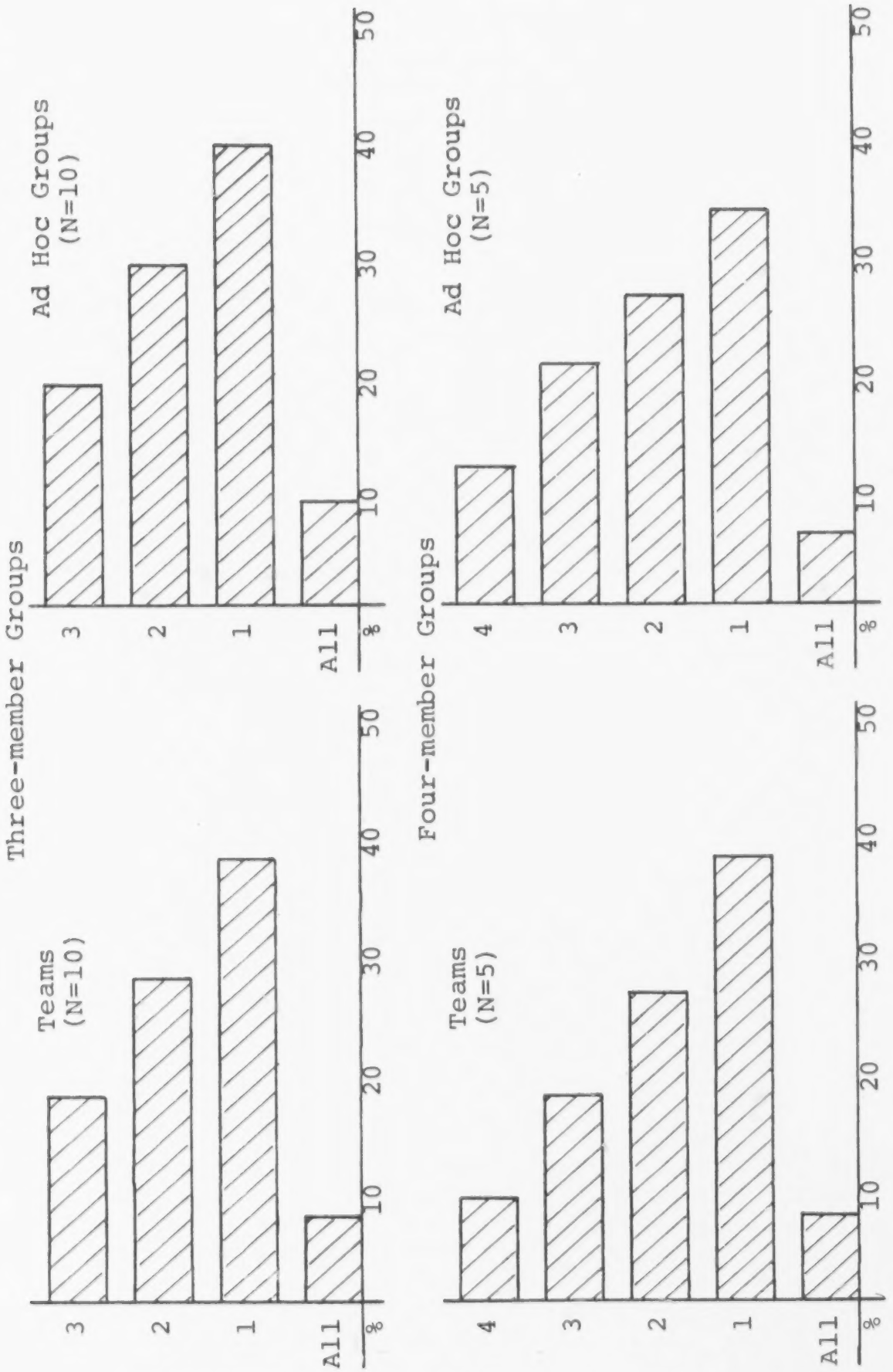
large groups than in small ones; however, more tension and antagonism seemed to appear in ad hoc groups than in teams and in teams with leaders than in teams without. These differences were reflected in the positive socio-emotional areas since more solidarity and agreeing behavior tended to occur in smaller groups than in larger ones and more agreeing behavior seemed to occur in teams than in ad hoc groups.

In general, this sample of teachers seemed to approach the given problems in the group situation by expressing individual opinions and/or analyses of the problems. Group members tended to agree and support each other more than they disagreed. Tension and antagonism were not very evident, especially in smaller groups and in those that had worked together for a period of time.

The interaction profiles may be somewhat helpful in understanding why experienced teams did not achieve better results than ad hoc groups with the DPT. As Figures 2 and 3 indicate, the two types of groups did not seem to differ in the amount of interaction devoted to the task, while the distribution of behavior within the task-related categories was also very similar. These profiles may also have some value in accounting for the failure of groups to excel over individuals on the DPT. There were relatively high amounts of agreeing behavior compared to disagreeing behavior in both types of groups. This may be evidence of initial consensus among group members regarding the task which Barnlund (1958) and Hall and Williams (1966) found to be detrimental to effective group problem solving.

Group Participation Patterns

The total number of interactions initiated per person per group was obtained from the observers' records. Group members were then ranked from those with the most to those with the least number of initiations. The average percent of participation was then found for all members who ranked first for each size and type of group, then for those who ranked second, third, etc. (see Figure 8). A visual inspection of these distributions shows that, regardless of size or type of group, the pattern of participation was not evenly distributed among group members. Furthermore, as group size increased, one or two persons tended to participate very little. It can also be seen that as group size increased, even the most frequent initiators tended to originate interactions more seldom than in smaller groups. The widest range of participation occurred in four-member teams (from a low of 10 percent to a high of 40 percent); whereas the range was narrowest (15 percent to 32 percent) in four-member ad hoc groups. Participation patterns for three-member teams and three-member ad hoc groups were almost identical. However, in four- and five-member teams there was a tendency for one person to dominate the discussion more than in ad hoc groups of the same size. This dominance pattern may


Distribution of Initiation of Interaction
Three-member Groups
Teams (N=10)
Ad Hoc Groups (N=10)
Four-member Groups
Teams (N=5)
Ad Hoc Groups (N=5)
4
3
2
1
All
%
10
20
30
40
50

Five-member Groups

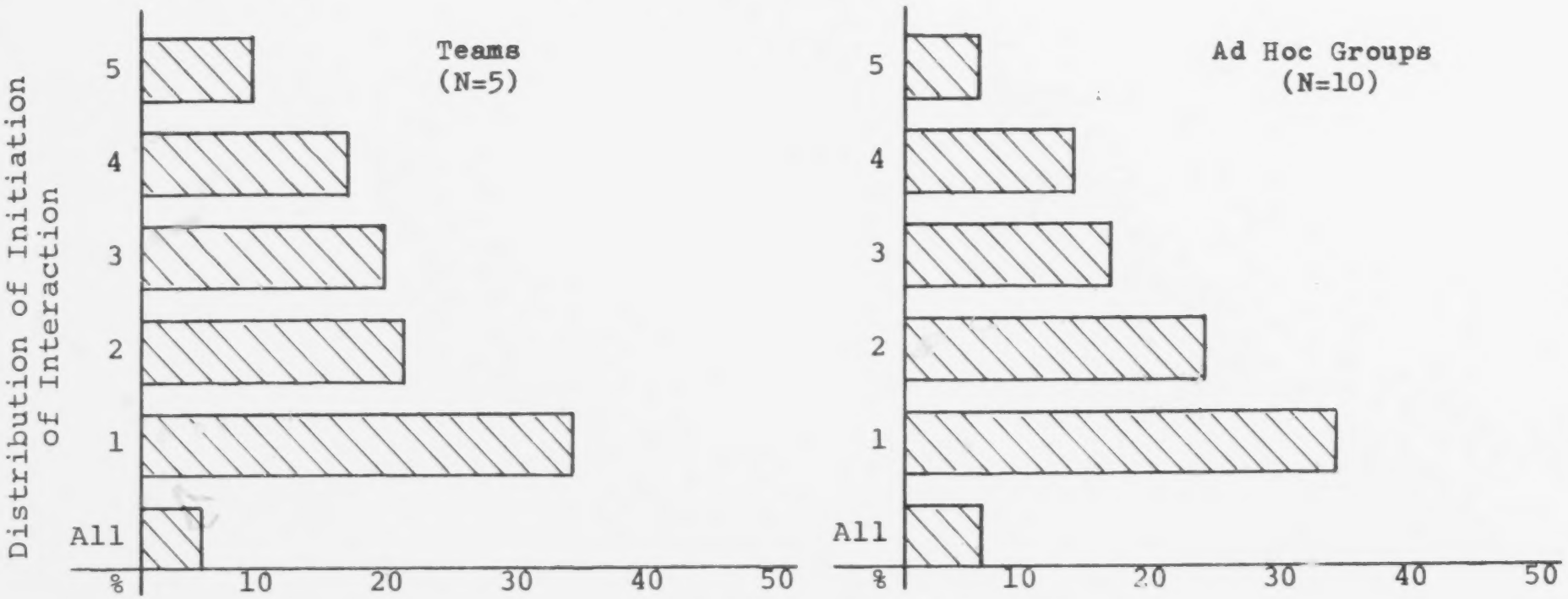


Fig. 8.--Distribution of Initiation of Interaction. Each bar represents a percentage of the total initiations of interaction within groups of that size and type. The number of persons represented by each bar varies with the type of group. In three-member teams, for example, Bar 1 stands for the average percentage of participation for those persons who ranked first in initiating interactions in that size and type of group. Bar 2 represents the average of those who were the next most frequent initiators, and so on.

have occurred because there was either a formal or an ascribed leader in the teams who attempted to influence the problem-solving process, whereas ad hoc groups tended to be leaderless.

Group participation patterns were examined in this study because several small-group studies (e.g., Davitz et al., 1952; Shaw, 1932; Thorndike, 1938) have shown that all group members do not participate to the same extent and that minority viewpoints are often not heard or incorporated into group solutions. Maier (1967) and Hoffman (1965) have identified the tendency of vocal members to talk down minority viewpoints as one of the inhibiting factors in group work. It appears that the members of both teams and ad hoc groups did not participate equally, especially in the larger groups. Although it is not possible to determine from the interaction data whether or not minority viewpoints were expressed and adequately discussed, it seems likely that where one person tended to dominate the discussion, all members' opinions may not have been heard. The unequal participation of group members may therefore help to account for the lack of differences between teams and ad hoc groups as well as between groups and individuals.

Two-Way Analysis of Variance on Group Interaction

Teams and ad hoc groups differed very little in the total amount of behavior that was categorized as task related; differences in means for each of the task-related categories making up total task behavior (4-9) were also very small. Furthermore, since the Bales Interaction Analysis is designed to record the quantity of interaction related to the task and not the quality of this interaction, further analysis of task behavior would contribute little to an understanding of the results obtained on the Decision Process Test. Therefore, no further analysis was made of task behavior. However, differences in means for the socio-emotional categories were large enough to suggest the need for further analysis. Therefore, a two-way analysis of variance for groups by size and type was run on these categories.

Results attributable to size of group. On Category 1 (Shows solidarity), a significant difference (.01) related to group size was found. Both teams and ad hoc groups displayed less of this behavior as the size of the group increased. There also was a tendency (.10) for the amount of behavior in Category 3 (Agrees) to decrease as the size of the group increased (see Table 13). In the negative socio-emotional behavior categories, there was an increase in disagreeing behavior as the size of the group increased. This difference was significant (.01), and there was also a trend (.10) for behavior in Category 11 (Shows tension) to increase as group size increased. In each case the change in behavior was most evident when the size of the group increased from four to five members (see group means, Table 14).

TABLE 13

Group Interaction: Two-Way Analysis of Variance for Groups by Size and Type

Variable	Source of Variation	Sum of Squares	df	F-ratio
Percentage of group behavior in Category 1 (Shows solidarity)	Group size	42.45	2	8.30**
	Group type	2.73	1	1.06
	Interaction	3.97	2	0.77
	Error	86.88	34	
	Total	136.03	39	
Percentage of behavior in Category 3 (Agrees)	Group size	45.49	2	2.91‡
	Group type	23.91	1	3.06*
	Interaction	4.88	2	0.31
	Error	265.67	34	
	Total	339.95	39	
Percentage of behavior in Category 10 (Disagrees)	Group size	34.37	2	13.58**
	Group type	0.06	1	0.05
	Interaction	3.17	2	1.21
	Error	43.04	34	
	Total	80.63	39	
Percentage of behavior in Category 11 (Shows tension)	Group size	7.35	2	2.41‡
	Group type	11.29	1	7.41*
	Interaction	2.67	2	0.88
	Error	51.83	34	
	Total	73.14	39	

*$p < .05$.

**$p < .01$.

‡$.05 \quad p < .10$.

Results attributable to type of group. One significant difference and one trend toward significance were noted due to type of group. Ad hoc groups displayed significantly more tension (.05), Category 11, than did teams. Teams also tended to show more agreeing behavior (.10), Category 3 (see Tables 13 and 14).

TABLE 14

Group Interaction: Means and Standard Deviations for Groups by Size and Type

Variable	Number in Group		Teaching Teams		Ad Hoc Groups	
Percentage of behavior in Category 1 (Shows solidarity)	3	Mean	6.59		6.52	
		SD		1.62		1.01
	4	Mean	6.90		5.38	
		SD		2.19		1.59
	5	Mean	4.10		4.04	
		SD		1.82		1.69
Percentage of behavior in Category 3 (Agrees)	3	Mean	12.49		11.72	
		SD		3.36		3.36
	4	Mean	12.64		10.20	
		SD		1.22		2.80
	5	Mean	10.34		8.66	
		SD		2.39		0.72
Percentage of behavior in Category 10 (Disagrees)	3	Mean	1.97		1.55	
		SD		1.15		0.95
	4	Mean	2.18		2.96	
		SD		0.87		0.64
	5	Mean	4.34		3.72	
		SD		1.25		1.75
Percentage of behavior in Category 11 (Shows tension)	3	Mean	1.95		2.41	
		SD		1.02		0.94
	4	Mean	1.56		3.22	
		SD		0.73		1.49
	5	Mean	2.60		3.84	
		SD		1.10		2.16

One-Way Analysis of Variance on Group Interaction

When the social interaction of teams with leaders was compared to that of teams without leaders, two significant differences and several trends were noted. Teams with leaders tended to display less total positive social behavior (Categories 1 to 3) than teams without leaders (.10) and significantly more total negative social behavior (Categories 10 to 12); Category 10 (Disagrees) was the primary source of the variance (.001). There was also a tendency for Category 11 (Tension) and Category 12 (Antagonism) to be

greater (.25) for teams with leaders compared to those without (see Table 15).

TABLE 15

Group Interaction: One-Way Analysis of Variance (N = 10)

Behavior Category		Teams with Leaders		Teams Without Leaders		F-ratio
1-3	Mean	19.70		22.46		
	SD		3.15		2.99	4.03
10	Mean	3.75		1.48		
	SD		1.22		0.46	30.19*
11	Mean	2.34		1.69		
	SD		0.97		0.97	2.24
12	Mean	0.53		0.21		
	SD		0.52		0.34	2.61
10-12	Mean	6.62		2.78		
	SD		2.02		0.78	31.54*

*$p < .001$.

These analyses seem to indicate that increasing group size may result in less positive and more negative socio-emotional behavior and that this may be more evident in newly formed groups than in experienced ones. Within experienced groups, this same tendency may be associated with the presence of a formal leader. However, these findings and the profiles made of group interaction were important to this study only insofar as they helped to account for results on the DPT. Therefore, a further analysis of interaction as it pertained to results on the DPT was made.

Significance of Interaction Data to Results on the Decision Process Test

Three interaction patterns that might account for the failure of teams and ad hoc groups to perform better than individuals on the two main variables of the DPT--rationality of decision behavior and quality of decisions--were identified. These were (1) a high percentage of negative social interaction, i.e., disagreement, tension, and antagonism; (2) a high percentage of behavior devoted to social interaction rather than concentration on the given task; and (3) an unequal distribution of participation, i.e., initiation of interaction dominated by one or two persons. Since no

significant differences were found between teams and ad hoc groups or between teams with leaders and teams without leaders when analyzing results on rationality and decision quality, the type of group was disregarded and the ten groups displaying the highest amounts of the behaviors noted above were compared by one-way analysis of variance to the ten groups exhibiting the lowest amounts of these behaviors. In this way it was hoped to find out whether any of these three patterns of interaction tended to interfere with the quality of problem solving by a group.

High vs low negative socio-emotional behavior. Ten groups (four teams and six ad hoc groups) had more than 7.5 percent of their behavior categorized as negative socio-emotional; ten other groups (six teams and four ad hoc groups) had less than 3.5 percent of their behavior categorized this way. The groups having high amounts of negative socio-emotional behavior were compared to the groups having low amounts of such behavior on the degree of rationality of their decisions, and the quality of their decisions. A significant difference (.01) was found on the degree of rationality. No difference was found on the quality of decisions.

As can be seen from Table 16, groups with low amounts of socio-emotional behavior had a mean of .70 on rationality of decision behavior (i.e., the correlation between actual and expected rankings of alternative teacher behaviors) in comparison with a mean of .47 for those groups with high amounts of socio-emotional behavior. The latter mean was very similar to the overall mean for individuals on this variable, which was .49 (see Table 4, variable 11). In this study, it appears that tension and antagonism of individuals within groups may have interfered with the group's ability to use prior thinking in arriving at a final ranking.

TABLE 16

Effects of High vs Low Socio-Emotional Behavior on Rationality and Quality of Decisions

Variable		High Negative		Low Negative		F-ratio
Rationality of Decision Behavior	Mean	.47		.70		10.00**
	SD		.20		.09	
Quality of Decision Reached	Mean	.57		.66		n.s.
	SD		.15		.13	

**p < .01.

High vs low task behavior. When the ten groups with more than 76.5 percent of their behavior categorized as task related were compared with the ten groups with less than 71.5 percent, there were no significant differences on either of the two variables for which comparisons were made. The range on task behavior for all groups was from 69 to 81 percent. Apparently, this range was not great enough to affect the task variables analyzed.

Even vs uneven patterns of participation. The ten groups with the most evenly distributed patterns of member initiation of interaction were compared to the ten groups in which the interaction was dominated by one or two group members. There was no significant difference for either variable. It appears that groups in which all members tended to participate equally were able to achieve no better or more rational decision, as measured by these variables, than were groups in which one or two members initiated most of the interaction.

Interobserver Reliability

The value of the interaction data in interpreting results on the Decision Process Test depends, in part, on how reliably the two observers were able to record the interactions that they saw. Therefore, tests of interobserver reliability were made. The Kendall rank order correlation coefficient (tau) was used to determine correlation between the two observers (Seigel, 1956, pp. 213-19). This test indicates whether or not each category was used to the same extent by each observer. A correlation of .80 or better was considered acceptable. Using this test, the mean correlation between observers for teams was .90 with a range from .79 to .95; only one correlation (.79) fell below the acceptable level. Five correlations were below the .80 level on ad hoc groups: one each at .69, .70, and .74, and two at .79.

The Kendall rank order correlation coefficient also was used to determine limits of interobserver reliability on the designation of the originators of interactions. The mean correlation between observers for teams was .96 with a range from .80 to 1.00; on ad hoc groups the mean was .91 with a range from .50 to 1.00. All correlations reached acceptable levels of reliability on teams, and seventeen out of twenty were satisfactory for ad hoc groups. Three correlations for ad hoc groups fell below the .80 level, one at .50 and two at .67. In each of these three cases, one observer appeared to attend more to one group member than did the other, or tended to use the O classification instead of identifying group members individually.

Although tape recordings were made of group interaction as it occurred, these were not used to try to improve the correlations between observers. Had the records for one group been reanalyzed, it would have been necessary to redo all forty groups. This did

not seem practical in terms of the cost of having typed transcriptions made for fifty or more hours of group interaction, especially in light of the difficulty in distinguishing between voices in all-female groups. Most of the differences were in recording ad hoc groups rather than teams. From the tapes, it was determined that the interaction in ad hoc groups tended to be more difficult to follow because group members tended to talk at the same time. It would have been just as difficult to interpret the speaker when typing the transcriptions as it was during the actual observation of behavior. Acceptable levels of correlation were achieved with the Kendall rank order correlation test except as noted above. Therefore, the limitations on reliability of interaction data need to be kept in mind in reviewing results.

SUMMARY OF FINDINGS

The Decision Process Test was used to compare four experimental conditions on twelve variables dealing with the process of decision making and five variables dealing with the quality of decision making according to four hypotheses: (1) groups will use more rational decision processes and make better decisions than individuals; (2) teams will use more rational processes and make better decisions than ad hoc groups; (3) teachers assigned to teams will differ from teachers assigned to self-contained classrooms on quality of decisions and on decision-making behavior; and (4) teams with leaders will differ from teams without leaders on quality of decisions and on decision-making behavior.

No support was found for the hypothesis that groups would use more rational decision processes and make higher-quality decisions than individuals. However, one main effect between groups and individuals and five interaction effects between one experimental condition and another were noted on decision-making variables. The main effect between groups and individuals involved the more frequent use by groups of the extreme ends of the desirability scale on Problem PA. This variable also accounted for two interaction effects with teams more extreme than individuals from teams on one problem and ad hoc groups more extreme than individuals from teams on another. Two interaction effects occurred on the variable of total working time with teams taking longer to work than individuals from teams on one problem and ad hoc groups taking longer than individuals from teams on another. The fifth interaction effect was found on the variable of probability direction with teams listing consequences they saw as more probable than did individuals from self-contained classrooms.

Generally, when subjects were required to think of consequences to the child, those that came to mind were ones the subjects felt had at least a 70 percent chance of occurring, were

likely to happen within a week, and were either very desirable or very undesirable.

Two interaction effects were found on decision quality, with the difference in each case favoring the group condition. The final ranking of alternatives by teams correlated more closely with the ranking by professors of elementary education than did the ranking by individuals from teams on Problem PA. The final ranking of ad hoc groups also correlated more closely with that of educational psychologists than did the ranking of individuals from self-contained classrooms.

Using one-way analysis of variance, no support was found for the hypothesis that teams would exhibit a more rational pattern of decision behavior and make higher-quality decisions than ad hoc groups. However, two significant differences were found supporting this hypothesis when two-way analysis of variance for groups by size and type was used. In comparison with ad hoc groups, teams ranked the alternatives for Problem PA more as they were expected to, given their prior evaluations of alternatives. One other significant difference was noted in decision behavior, which apparently was due to group size. Teams were found to become more positive in evaluating the desirability of the perceived consequences of the alternative teacher actions as the size of the team increased, whereas ad hoc groups became more negative.

No support was found for the hypothesis that team teachers would differ from self-contained-classroom teachers in decision behavior. One significant difference was found in decision quality, but this was dismissed as random variation. Furthermore, there were no significant differences to support the hypothesis that teams with leaders would differ from teams without leaders.

Group interaction data were collected by two observers in order to determine whether the manner in which group members interacted with each other tended to lower the quality of group solutions to problems on the DPT since this might account for the failure to find support for the first two hypotheses.

Pooled interaction profiles of the categorization of group behavior (see pp. 33-38) revealed that approximately 74 percent of the interaction was related to the task and 26 percent was socio-emotional. Task behavior tended to consist mainly of giving opinions, making analyses, obtaining information, or clarifying. Teams and ad hoc groups did not differ significantly in the amount of interaction involved in carrying out the task.

Socio-emotional behavior was generally positive rather than negative. However, two-way analysis of variance revealed some significant differences in social behavior that were related to the size and type of group. Both teams and ad hoc groups showed

significantly more solidarity (.01) in small groups than in large ones and more disagreeing behavior (.01) in large groups than in small ones. Teams with leaders showed more total negative behavior (.001) and more disagreeing behavior (.001) than teams without leaders. Ad hoc groups displayed more tension (.05) than did teams.

Groups (either teams or ad hoc groups) displaying extreme interaction patterns were compared on the variables of rationality and quality of decisions by one-way analysis of variance. This analysis revealed that groups with low amounts of negative social behavior were more consistent in ranking alternative teacher actions in line with their prior thinking than were groups with high amounts of negative social behavior (.01). However, the type of social behavior did not seem to affect the quality of decisions. In addition, it could not be shown that either higher amounts of task-related behavior or equal participation of all group members contributed to higher scores on these same variables.

DISCUSSION

Those who recommend teaching teams as a form of elementary school organization claim that teachers working as groups will make better decisions for students. The basis for this claim is a general conclusion regarding the superiority of group performance over individual performance that has been derived mainly from small-group research. Most of this research has been done with ad hoc rather than experienced groups, using tasks unrelated to those involved in teaching. Principles derived in this way are not necessarily applicable to teaching.

This study was designed to test the assumption of group superiority in decision making under conditions that would allow the conclusions reached to be applied to actual teaching practice. These conditions were as follows: ad hoc and established teams were used; teachers from teams and from self-contained classrooms were assigned to both the group and the individual condition; the experimental task was a teaching problem with which the subjects were familiar and at least moderately concerned; and comparative data were available from a previous study (Brim et al., 1962) that had used an instrument very similar to the Decision Process Test used in this study.

However, three possible limitations need to be kept in mind in applying the results of this study to teaching practice: the sampling procedures used; the procedures for administering the DPT; and the validity of the DPT.

Sampling Procedures

Data were collected on the age, sex, experience, and training of subjects; these data could not be consistently associated with results on the DPT. Although subjects were comparable on these factors (see Table 2), because sampling procedures were not completely random it cannot be assumed that they did not differ on other important characteristics. Therefore, it seems likely that observed variations among subjects may have been dependent upon such factors as intelligence and teaching skills of which there were no measures.

Participating school districts were not chosen randomly either. Only medium-sized, middle-class, suburban school districts were included in this study. All of these districts were experimenting with team teaching despite the fact that this innovation was probably costing more than a self-contained organization. Such districts may not be typical of school districts across the nation and therefore the results may not be widely generalizable beyond the type of district used in the study.

Procedures for Administering the Test

Teachers in the individual condition were assembled within districts as much as possible so that the DPT could be administered to several of them at a time. The results of some small-group research suggest that individuals may work harder on a task when they are in the company of others, even though there is no interaction. Therefore, it is possible that differences between groups and individuals were affected by this.

Validity of the Decision Process Test

The content validity of the Decision Process Test was assessed (see pp. 26-28), but no attempt was made to establish the adequacy of a rational decision model for studying the actual decision behavior of teachers either in groups or as individuals. The subjects in this study, and in the study by Brim et al. as well, did not appear to consider a full range of consequences, i.e. they concentrated on those they saw as quite likely to happen, very desirable or very undesirable, and likely to happen in the near future.

An analysis of the consequences listed in this study also revealed that subjects did not always respond in terms of specific consequences. For example, frequent types of responses were like one or more of the following: (1) the stated action would have "no effect" on the child's behavior; (2) the action would make the child's behavior "better"; (3) the action would make the child's behavior "worse." All three types of responses were given for each alternative by some subjects. The space in which consequences

were to be listed was also frequently used to qualify the given teacher action rather than to list consequences of it as stated.

The above results suggest that teachers (and possibly parents) may be more accustomed to using a trial-and-error approach to modifying a child's behavior than to analyzing the possible alternatives open to them before they take act on. If this is true, a rational decision model would not describe their actual decision behavior in a real situation.

The decision model also assumes that adults will choose the alternative offering maximum value to the child. Both the teachers in this study and the parents in Brim's study tended to omit consequences that were not likely to occur, that were not as severe for the child, and that would have long-range effects. They may have omitted the first two on this instrument to make the task more manageable. However, omitting long-range consequences from consideration may be evidence of poor decision behavior or evidence that subjects were choosing among alternatives in a way that would maximize value for themselves to some extent. They may have been concerned with stopping the undesirable behavior immediately without thinking of the possible long-range consequences of their actions on the child.

Tests were run on the data to determine whether the expected utility score on each alternative (P x D) would predict first choices of subjects at a better-than-chance level. This score did prove to be a good predictor in all except one experimental condition (see Table 17).

Although the above analysis lends some support to the rational decision model on which the DPT was based, it does not appear that the subjects used a rational process to eliminate alternatives consistently, even when forced to consider consequences first. The pooled mean correlation between the expected ranking of alternatives and the actual ranking of these across experimental conditions and problems was 0.54 with a range from -0.33 to 1.00. Nine out of 160 correlations fell at 0.00 or below, while seven were 0.90 or better. It was possible to predict first choices only about 56 percent of the time.

If teachers did not use prior evaluations of consequences in arriving at a ranking among alternatives, what actually did influence their choices among alternatives? One possible influence might have been existing norms for handling this type of behavior. To test this, an average ranking for the alternatives of each problem was computed for the original sample of teachers who suggested the alternatives used on the DPT. This was compared with the average ranking by subjects in this study using Kendall's rank order correlation test. On Problem PA these rankings correlated 1.00, while on Problem IB the correlation was 0.60. It seems

TABLE 17

Number of Correct and Incorrect Predictions
Based on Expected Utility Scores (P x D)

Problem	Group Condition				Individual Condition			
	Teams (N=20)		Ad Hoc Groups (N=20)		Team Teachers (N=20)		S-C-C Teachers (N=19)	
	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong
IB	11*	9	12**	8	12**	8	7	12
PA	13**	7	11*	9	12**	8	10*	9

*p < .05.

**p < .01.

possible, therefore, that subjects might have ranked these alternatives in the same way even if they had not considered consequences first. It would have been very helpful to have had a ranking of alternatives both before and after consideration of consequences to determine whether the evaluations would have changed the final ranking.

CONCLUSIONS

Looking at the results of the study with the above limitations in mind, there is very little evidence to support the belief that groups of teachers make better decisions than individual teachers. The groups and the individuals tended to behave very much alike and to reach decisions that were, according to the experts, qualitatively the same. The only main effect found did not appear to affect the quality of the decision.

Because groups were more extreme in the way they rated the consequences of teacher behavior, it appears that discussing these consequences with others had the effect of intensifying feelings. Brim et al. found the direction of desirability to be more positive for groups than for individuals and found only one interaction effect (.05) on the extremity of desirability; however, they used

only established two-member groups. This study included ad hoc groups, which tended to become more negative as they grew larger (see Table 14). This tendency may have offset the more positive evaluations of teams (established groups) and made groups and individuals appear to be similar. Extremity in either direction also increased with group size, in this study. The use of two-member groups by Brim et al. may account for the fact that they found few differences, whereas in this study a main effect was noted.

Extremity owing to group discussion may be evidence of one of the liabilities of group problem solving that Maier (1967) has identified. Social pressure to reach a decision and achieve consensus can silence minority opinions, regardless of their worth. If extremity of thinking is evidence of this type of pressure, then it is possible that group decisions might have been of better quality had the groups been more similar to the individuals on this variable.

Brim et al. found differences between groups and individuals on time direction and extremity. Similar trends were found in this study, but the differences were not significant. The type of student problem may account for this result. Brim et al. found differences on the problems of masturbation and stealing, and similar trends for obedience and homework. In this study, differences between groups and individuals were slightly greater for time direction and extremity on Problem PA (see Table 4, variables 9-10). It is possible that the more emotionally laden a problem seems to be, the more one tends to seek immediate solutions without regard to long-term effects. If this is true, group discussion might be expected to intensify such feelings.

One finding in the study by Brim et al. led to the hypothesis that groups would be more rational in the way they ranked alternatives after having considered the possible consequences. This was not the case. It is possible that the factor of group size affected the rationality of the groups in this study. The two-way analysis of groups by size and type, for example, revealed that both teams and ad hoc groups tended to become less rational when the group reached five members (see Table 7).

The tendency of the subjects in this study to think of probable consequences and those likely to happen soon was noted also by Brim et al. This may be a function of the way the subjects saw the task in both studies. If this tendency is typical of actual problem-solving behavior, however, it may be something that teachers need to overcome. There is need for descriptive models of how teachers actually solve problems.

Very little evidence was found that ad hoc groups and established teams differed in their approaches to the two problems except that the interpersonal relations of team members were

somewhat more positive. There were two significant differences on Problem PA that favored teams over ad hoc groups on the ranking of alternatives in line with the prior evaluations of them (see Table 8). These may have occurred because team members seemed to be able to disagree with less tension and antagonism than ad hoc groups and may have been able, therefore, to use their prior evaluations to a greater extent. Hall and Williams (1966) found that established groups with a high initial potential for disagreement apparently used their disagreements more constructively and creatively in arriving at decisions than did ad hoc groups having a similar potential for disagreement. It is possible that the business executives in Hall and Williams's study may have had training in group decision making that the teachers in this study did not have, and were for that reason able to use their disagreements to better advantage.

There were no significant main effects between groups and individuals on total decision time, although there were two interaction effects (.05) between experimental conditions. What seems more important, however, is a consideration of the total number of man hours spent on the problems in the group versus the individual conditions. When the actual time spent by each group is multiplied by the number of group members and the mean number of man hours computed, groups are found to have spent a mean of 4.8 man hours on the two problems compared to a mean of 1.0 man hours for individuals. In view of the results on other variables, these extra man hours did not yield great returns.

No conclusions can be drawn concerning team leadership, but many questions were raised by the study. Smaller teams without appointed leaders tended to get along better and to handle a decision-making task more efficiently. Obviously more research in the area of team organization is needed. Judging from the training and the behavior of the elementary teachers who participated in this study, however, the hierarchical model for team teaching advocated by some educators (e.g., Anderson, 1966; in Shaplin & Olds, 1964) does not appear to be as effective as the less structured cooperative team.

Another question concerning the hierarchical model can be raised from the findings in this study regarding group size. A team that accommodates different levels of teaching responsibility (e.g., team leader, master teacher[s], teachers, and aides) tends to be somewhat larger than the cooperative team. Variations in team size were found among school districts: some teams had as few as two members, other as many as seven or eight. Only three-, four-, and five-member teams were selected for this study. Changes in efficiency and interpersonal relations apparently began to occur when the group grew from four to five members. Greater effectiveness has often been associated with smaller groups in small-group studies (McGrath, 1966), and this factor needs to be

examined in a school setting before decisions can be made regarding the optimum size for a teaching team.

Team teaching may be very valuable for reasons other than those explored in this study. However, these findings do not appear to warrant the assumption that teams will necessarily make better decisions than will individual classroom teachers. Neither were there big differences between teams and ad hoc groups, except perhaps in the area of interpersonal relations. Length of experience alone may not be enough to enable teams to improve their problem-solving skills. There may be a need for special team training to overcome the liabilities of group problem solving and to realize the potential of the group. Finally, much more research is needed to identify the most effective size and type of organization for a teaching team.

REFERENCES

Anderson, Robert H. Teaching in a World of Change. New York: Harcourt, Brace and World, 1966.

Bales, Robert F. Interaction Process Analysis. Cambridge, Mass.: Addison-Wesley, 1950.

Barnlund, Dean. "A Comparative Study of Individual, Majority, and Group Judgment," Journal of Abnormal Social Psychology, 58 (1959), 55-60.

Brim, Orville G., Jr., et al. Personality and Decision Processes. Stanford, Calif.: Stanford University Press, 1962.

Collins, Barry E., & Guetzkow, Harold. A Social Psychology of Group Processes for Decision-Making. New York: John Wiley & Sons, 1964.

Davitz, Joel; Lorge, Irving; Herrold, Kenneth; & Fox, David. "The Evaluation of Group versus Individual Decisions," American Psychologist, 7 (1952), 313. Abstract of a paper read at the Sixtieth Annual Meeting of the American Psychological Association.

Guenther, William C. Analysis of Variance. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Hall, Jay, & Williams, Martha. "A Comparison of Decision-Making Performances in Established and Ad Hoc Groups," Journal of Personality and Social Psychology, 3 (1966), 214-22.

Heathers, Glen. "School Organization: Nongrading, Dual Progress, and Team Teaching," in John I. Goodlad, ed., The Changing American School. The Sixty-Fifth Yearbook of the National Society for the Study of Education, Part II. Chicago: The University of Chicago Press, 1966.

Hillson, Maurie. Change and Innovation in Elementary School Organization: Selected Readings. New York: Holt, Rinehart and Winston, 1965.

Hoffman, L. Richard. "Group Problem Solving," in Leonard Berkowitz, ed., Advances in Experimental Social Psychology, II. New York: Academic Press, 1965. Pp. 99-175.

Knox, Donald M. "An Experimental Study of the Effect of a Team Teaching Program upon Certain Selected Variables (Achievement-Anxiety-Social Relations)." Unpub. Ph.D. dissertation, Western Reserve University, 1965, in Dissertation Abstracts, 27, No. 2 (1966), 416-17.

Lambert, P., et al. Classroom Interaction, Pupil Achievement, and Adjustment in Team Teaching as Compared with the Self-Contained Classroom. Cooperative Research Project No. 1391. Madison: University of Wisconsin, 1964.

Lorge, Irving, et al. "A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance, 1920-1957," Psychological Bulletin, 55, No. 6 (1958), 337-72.

Lorge, Irving; Tuckman, Jacob; Aikman, Louis; Spiegel, Joseph; & Moss, Gilda. "Solutions by Teams and by Individuals to a Field Problem at Different Levels of Reality," Journal of Educational Psychology, 46 (1955), 17-24.

McGrath, Joseph E., & Altman, Irwin. Small Group Research. New York: Holt, Rinehart and Winston, 1966.

Maier, Norman R. F. Problem-Solving Discussions and Conferences, Leadership Methods and Skills. New York: McGraw-Hill, 1963.

Maier, Norman R. F. "Assets and Liabilities in Group Problem Solving: The Need for an Integrative Function," Psychological Review, 74, No. 4 (1967), 239-49.

Shaplin, Judson T., & Olds, Henry F., Jr., eds. Team Teaching. New York: Harper & Row, 1964.

Shaw, Marjorie E. "A Comparison of Individuals and Small Groups in the Rational Solution of Complex Problems," in Eleanor E. Maccoby, Theodore M. Newcomb, and Eugene L. Hartley, eds., Readings in Social Psychology. New York: Holt, Rinehart and Winston, 1958. Pp. 564-75.

Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Simon, Herbert A. Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations. New York: The Macmillan Company, 1961.

Taylor, Donald W. "Decision Making and Problem Solving," in James G. March, ed., Handbook of Organizations. Chicago: Rand McNally, 1965. Pp. 48-86.

Thorndike, Robert L. "The Effect of Discussion upon the Correctness of Group Decisions When the Factor of Majority Influence is Allowed For," The Journal of Social Psychology, 9 (1938), 343-62. (a)

Thorndike, Robert L. "On What Type of Task Will a Group Do Well?," Journal of Abnormal Social Psychology, 33 (1938), 408-12. (b)

APPENDIXES

APPENDIX A: DECISION PROCESS TEST

Problem (PA)

Behavior: Johnny Moore seems to be very quick tempered. If he gets angry with another child, either in the room or on the playground, he is apt to hit or kick the child. He frequently gets involved in fights on the playground.

It is very important to know how familiar and concerned you are generally with the problem of aggressive behavior. Therefore, please circle one number on each of the following scales:

This type of behavior occurs in my (our) classroom:

A Great Deal	Quite a Bit	Some	Not Much	Not at All
1	2	3	4	5

I (we) consider this behavior to be:

Extremely Serious	Very Serious	Moderately Serious	Not Very Serious	Unimportant
1	2	3	4	5

Consequences of Teacher Behavior (Problem PA)

This is an action the teacher might take:	Isolate the child from the children with whom he has trouble.			
If a teacher took this action, what are some possible results? You may write in as many as five. Write only one result in each box.				
What is the probability that the result you have written would actually happen?	__Highly probable .95 __Probable .75 __Half and half .50 __Improbable .25 __Very improbable .05	__Highly probable .95 __Probable .75 __Half and half .50 __Improbable .25 __Very improbable .05	__Highly probable .95 __Probable .75 __Half and half .50 __Improbable .25 __Very improbable .05	__Highly probable .95 __Probable .75 __Half and half .50 __Improbable .25 __Very improbable .05
How much would you want this result to happen?	__Strongly desire __Desire __Don't care either way __Do not desire __Strongly do not desire	__Strongly desire __Desire __Don't care either way __Do not desire __Strongly do not desire	__Strongly desire __Desire __Don't care either way __Do not desire __Strongly do not desire	__Strongly desire __Desire __Don't care either way __Do not desire __Strongly do not desire
How soon would you expect this result to happen?	__Within a week __A week to a month __1 to 6 months __6 months to a year __1 year or more	__Within a week __A week to a month __1 to 6 months __6 months to a year __1 year or more	__Within a week __A week to a month __1 to 6 months __6 months to a year __1 year or more	__Within a week __A week to a month __1 to 6 months __6 months to a year __1 year or more

Check the items of information from the cumulative folder that you considered.

__Photograph
__Psychological referral
__Work samples
__Report card
__Absence record
__Health record
__IQ score
__Parental information
__None of these
__Achievement test scores
__Sibling information

Note: The instrument actually used allowed for five answers.

Choice of Actions

1. Now suppose that a teacher could take only one of these actions. Read over the actions and rank them below from the best to the worst single action a teacher might take by numbering them from 1 to 6. You may feel that none is any good or worth doing, but rank them anyway according to which is the best of the group, second best, and so on.

Course of Action	Rank Order
Isolate the child from the children with whom he has trouble.	_____
When Johnny is not present, help the other children to understand his problem and let them suggest ways to help him.	_____
Send him to the principal whenever this behavior occurs.	_____
Let the child express his aggressive feelings in more acceptable ways.	_____
Exclude him from the class and/or the playground whenever this behavior occurs.	_____
Let the other children treat him in the same way so he'll find out how it feels.	_____

2. Now write in the rank order number or numbers of the action(s) from among the six that you would actually take in this type of situation.

Write the number(s) here: _____

I (we) can't decide. _____

I (we) wouldn't do any of these. _____

Instead, I (we) would do the following: (Write in the action you would take. You may also use this space to suggest an action you might take in addition to the ones given above which you feel are appropriate.)

_____

_____

3. If, in Number 2 above, you couldn't decide what to do, was it because you did not have enough information about the problem?

Yes ____ No ____

If your answer is yes, what additional information would you have liked to have?

_____

_____

If your answer is no, can you give a reason why you were unable to decide?

_____

_____

APPENDIX B: EXPERT RATING SHEET

Problem (PA)

Johnny Moore seems to be very quick tempered. If he gets angry with another child, either in the classroom or on the playground, he is apt to hit or kick the child. He frequently gets involved in fights on the playground.

The following are courses of action which teachers sometimes take in handling the aggressive behavior of children. Assuming that a teacher could take only one of these actions, please rank order them from the best to the worst single action by numbering from 1 to 6. Given the information in the cumulative folders about Johnny Moore, you may feel that none of the stated actions is good. Please rank them anyway according to which is best of the group, second best, and so on. You are given space to rank them twice if you feel they should be ranked differently for a second-grade child than for a fifth-grade child; if not, rank them just once.

Teacher Action	Rank Order (Second Grade)	Rank Order (Fifth Grade)
Isolate the child from the children with whom he has trouble.	____	____
When Johnny is not present, help the other children to understand his problem and let them suggest ways to help him.	____	____
Send him to the principal whenever this behavior occurs.	____	____
Exclude him from the class and/or playground whenever this behavior occurs (i.e., make him sit in the office or send him home).	____	____
Let the other children treat him in the same way so he will find out how it feels (e.g., hit him back).	____	____

The following are also possible ways of handling Johnny's aggressive behavior. Please read through the list and rate them as follows:

1. Place a "+" beside any of those actions listed below which you feel are better than those you ranked on the previous page.
2. Place a "-" beside any of those actions listed below which you feel are poorer than those you ranked on the previous page.
3. You do not need to mark those actions which you feel are generally equivalent in quality to any of those you ranked on the previous page.

Rating	Teacher Action
____	Seek professional counseling help within the district.
____	Request a complete physical examination.
____	Request a neurological examination.
____	Transfer him to a "special" class.
____	Assign him to a teacher who is understanding of his problem.
____	Hold a parent-teacher conference.
____	Hold a parent-teacher-child conference.
____	Make a complete case study of the child.
____	Teach a unit of social problems.
____	Isolate the child from others with constructive things to do, such as a "job" in the library.
____	Help build rapport between the child and the principal.
____	Let him make up his own standards for behaviors and, if he cannot follow these, help him to revise them.
____	Give him responsibilities in the room or around the school.
____	Reward him for his good behavior.
____	Help him find new friends.
____	Let him decide what his own punishment will be.
____	Assign him a "buddy" who will tolerate his behavior.
____	Try to reach him through his friends.
____	Help him to understand why he does these things and to think through better ways of reacting.
____	Counsel the child; reason with him.
____	Know the child; establish a good rapport with him.

TABLE B-1

Average Rank Given to Alternatives by Subgroups of Experts and Subjects

Problem	Alternatives	Elementary Educators		Educational Psychologists		Child Growth & Development Specialists		Ss in the Study	
		Primary	Intermediate	Primary	Intermediate	Primary	Intermediate	Primary	Intermediate
IB	Seat child near front of room	4	4	2.5	2.5	1.5	1	2	2
	Have child repeat teacher's instructions	5	5	5	5	5	5	5	5
	Ignore the behavior	6	6	6	6	6	6	6	6
	Teach lessons on listening	3	3	2.5	2.5	4	3	4	4
	Reward the child when he listens	1	2	1	1	1.5	2	1	1
	Explain importance of listening	2	1	4	4	3	4	3	3
PA	Isolate from those with whom he has trouble	3	2	3.5	3.5	1.5	1	3	3
	Help other children to understand	2	3	2	2	3.5	4	2	2
	Send him to principal	5.5	5	5	5	6	6	5	5
	Let him express aggressiveness in other ways	1	1	1	1	1.5	2	1	1
	Exclude him when this happens	5.5	6	3.5	3.5	3.5	3	4	4
	Let other children treat him in the same way	4	4	6	6	5	5	6	6

TABLE B-2

Correlations[a] Between Average Rank Given Alternatives for Subgroups of Experts and Total Group of Experts

Alternatives of Problem	Experts	Experts: Educational Psychologists		Experts: Child Growth & Development Specialists		Experts: Total	
		Primary	Intermediate	Primary	Intermediate	Primary	Intermediate
IB	Elementary Educators	.67	.53	.67	.33	.73	.67
	Educational Psychologists			.80	.80	.80	.80
	Child Growth and Development Specialists					.93	.73
PA	Elementary Educators	.53	.40	.47	.47	.60	.73
	Educational Psychologists			.47	.40	.67	.67
	Child Growth and Development Specialists					.87	.73

[a]Using Kendall's Rank Order Correlation Coefficient, Tau.